PL-TR-92-2101

CONVOLUTION ALGEBRA FOR FLUID MODES WITH FINITE ENERGY

G. v.H. SANDRI
C. KONSTANTOPOULOS

Boston University College of Engineering
and Center for Space Physics
Boston, MA 02215

April 1992

Scientific Report No. 3

Approved for public release; distribution unlimited;

PHILLIPS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AIR FORCE BASE, MASSACHUSETTS 01731-5000

"This technical report has been reviewed and is approved for publication"

ROBERT R. BELAND
CONTRACT MANAGER

DONALD E. BEDO
BRANCH CHIEF

ROGER VAN TASSEL
DIVISION DIRECTOR

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify PL/TSI, Hanscom AFB, MA 01731-5000. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1992	Scientific No. 3

4. TITLE AND SUBTITLE
Convolution Algebra for Fluid Modes With Finite Energy

5. FUNDING NUMBERS
PE 61101F
PR 7670 TA15 WU AR

Contract F19628-88-K-0017

6. AUTHOR(S)
G. v.H. Sandri
C. Konstantopoulos

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Boston University College of Engineering
and Center for Space Physics
Boston, MA. 02215

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Phillips Laboratory
Hanscom AFB, MA 01731-5000
Contract Manager: Robert Beland/OPA

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
PL-TR-92-2101

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release;
Distribution unlimited;

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A new set of parametric wavelets called *Hermite-Rodriguez* wavelets, and a new family of highly singular functions called hyperdistributions are introduced. Their introduction is motivated by the search for a consistent solution to the inverse problem in signals and systems analysis: the evaluation of the initial condition –or input– to a system , given its final condition –or output– and its impulse response. We construct a novel mathematical framework for the search for a solution to the inverse problem, and apply it to a practical case: the inverse problem for blurring systems. Blurring systems are very common: they include telescopes and other optical systems, as well as satellite telecommunications, where the waveforms that transmit the relevant information have to travel long distances, and are partly scattered by molecules and ions in their path from the source to the observer. As a result, the waveforms that reach the observer are blurred. The goal is to recover the original waveforms. The mathematical equivalent for the solution to the inverse problem for blurring systems –in the case of gaussian blurring– is given. It consists in solving the antidiffusion equation, which is equivalent to solving the diffusion equation backwards in time, to determine the initial condition of the diffusive process, given the final condition. We apply this solution of the inverse problem to two practical cases: the reconstruction of blurred signals, and of blurred simulated astronomical images. We then extend these results to the more general case of non-gaussian blur.

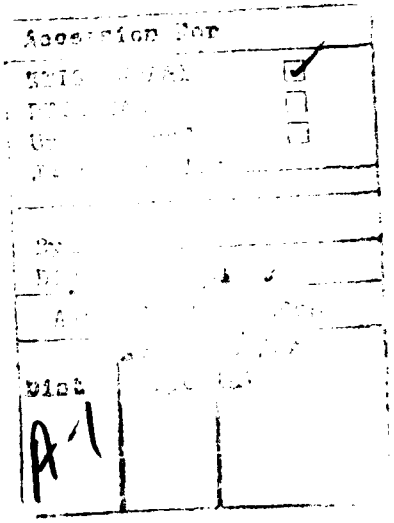
14. SUBJECT TERMS
Wavelet Analysis, Inverse Problems, Diffusion, Antidiffusion,
Deconvolution, Signal Enhancement, Image Restoration,
Hermite Polynomials, Singular Functions, Hyperdistributions

15. NUMBER OF PAGES
90

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

Contents



List of Figures

SUMMARY

Our analysis of the free sheared atmosphere is based on the hierarchy of equations satisfied by the modal coefficients (strengths). This hierarchy is obtained by substituting a series expansion of the fluid variables into the governing equations. Our analysis is particularly focused on the stability of the Taylor-Dyson atmosphere, which is designed to be a model of the free sheared atmosphere. The construction of the hierarchy is however independent of the model that one chooses albeit the more complicated the model, the more complicated the hierarchy. The classical approach considers Fourier series and the prototype is the successful Salzman-Lorenz hierarchy for the Benard problem, i.e. a thin layer of fluid heated from below. This problem is designed to be a model of the lower atmosphere. The application to the free sheared atmosphere modeled by the famous Taylor-Dyson equations of Fourier techniques has yielded the puzzling result that no unstable modes appear in linear order. The origin of the Richardson number criterion thus remains obscure. With the advent of wavelet analysis, we have deemed it useful (actually imperative) to perform the analysis, not on the basis of the Fourier modes which have infinite energy and essentially unbounded support, but rather on the basis of wavelets which are designed to have finite energy and essentially compact support. It is intuitively clear that such refined modal analysis should allow for a much more satisfactory physical interpretation of results. This point of view is further strengthened by the current trend in experimental work that emphasizes disturbances with finite spatial and temporal extents. At Boston University, we have developed a full form of wavelet expansion which has the advantage over more conventional procedures of possessing a complete convolution algebra.

The possibility of operating with a convolution algebra has so far been a prerogative of the Fourier analysis. Its major use in the present context is the calculation of the effect of forcing on a particular hierarchical level. Convolution algebra permits an immediate expression for the result of a forcing term on a variable that satisfies a given differential equation. The

mathematical tools needed to obtain a convolution algebra for modes with finite energy are rather extensive and not always elementary. To help our thinking, and also to help introduce interested scientist to a rather difficult area, we put together the most essential parts of the analysis and we present them in this report. It is worth mentioning at this point that very fruitful applications have been made of our wavelet analysis in biomedical areas. In particular the analysis of M waves in muscle response has been analysed very successfully with the wavelets presented here. While our main interest is the development of the chaos dynamics of the free sheared atmosphere, the Taylor-Dyson model in particular, we deem the mathematical background essential and also of general interest.

1 Signals and Systems and the Inverse Problem

1.1 Introduction

A majority of systems, natural and man-made, are *linear* and *time-invariant* (LTI)[1]. By linear, it is meant that if u_1 and u_2 are two inputs to a physical system, with respective outputs y_1 and y_2, then the output to the input $\alpha u_1 + \beta u_2$ is $\alpha y_1 + \beta y_2$, where α and β are two real numbers. By time-invariant, it is meant that a time-shift in the input signal causes an equal time-shift in the output signal. In other words, if the input $u(t)$ produces the output $y(t)$, then the input $u(t-\tau)$ produces the output $y(t-\tau)$. These properties are depicted in figure 1. It can be proven that the output of any linear time-invariant system can be modelled by the *convolution* of the input signal with the *impulse response* of that system. The impulse response (or point-spread function PSF) of a system is the output when the input to the system is a Dirac delta function (e.g. for electrical systems, a very high, very short-duration electrical pulse). The proof is as follows. Let $\mathcal{L}$ denote a discrete LTI system. Let a discrete Dirac delta $\delta[n]$ be the input to the system. By definition, the discrete impulse response $h_0[n]$ is the output of:

$$\delta[n] \quad \longrightarrow \boxed{\mathcal{L}} \longrightarrow \quad h_0[n] \tag{1}$$

Because of the property of shift-invariance,

$$\delta[n-k] \quad \longrightarrow \boxed{\mathcal{L}} \longrightarrow \quad h_k[n] = h_0[n-k] \tag{2}$$

Furthermore, because of the property of linearity,

$$\sum_k x[k]\delta[n-k] \quad \longrightarrow \boxed{\mathcal{L}} \longrightarrow \quad \sum_k x[k]h_k[n] \tag{3}$$

It can be proven that any discrete signal $x[n]$ can be written in the form $x[n] = \sum_k x[k]\delta[n-k]$. This property is commonly referred to as the *discrete sampling property*[1], or, more technically, as *Dirac's identity*. Combining Eqs.(2) and (3), we obtain

$$x[n] = \sum_k x[k]\delta[n-k] \quad \longrightarrow \boxed{\mathcal{L}} \longrightarrow \quad y[n] = \sum_k x[k]h_0[n-k] \tag{4}$$

In the case of continuous-time LTI systems, Eq.(4) is written as:

$$x(t) = \int_{-\infty}^{+\infty} x(\tau)\delta(t-\tau)\,d\tau \quad \longrightarrow \boxed{\mathcal{L}} \longrightarrow \quad y(t) = \int_{-\infty}^{+\infty} x(\tau)h_0(t-\tau)\,d\tau = x(t) * h_0(t) \tag{5}$$

where $h_0(t)$ is the impulse response of $\mathcal{L}$.

In many instances, the physicist, the engineer, or the mathematician is asked to determine the initial condition (or input) to a system, when the final condition (or

output) is known. This problem is known as the *inverse problem*. In mathematical terminology, knowledge of the system implies knowledge of its impulse response; the solution of the inverse problem implies knowledge of the output, and *convolution inverse* to its impulse response. This last statement is proven below. If u represents the input to the system, y the output, and h its impulse response, then by definition:

$$y(x) = h(x) * u(x) = \int_{-\infty}^{+\infty} h(x')u(x - x')\,dx' \tag{6}$$

If the impulse response h has a convolution inverse $Inv[h]$, this implies by definition that the convolution of h with $Inv[h]$ yields the Dirac delta function:

$$h(x) * Inv[h](x) = \delta(x) \tag{7}$$

That is because the Dirac delta is the unit element for the operation of convolution, much in the same way the number 0 is the unit element for integer addition:

$$\begin{array}{rcll} f(x) * \delta(x) & = & f(x) & \text{for any function f(x)} \\ n + 0 & = & n & \text{for any integer n} \end{array} \tag{8}$$

If y and h are known, then the initial condition u can be determined in the following way: convolve Eq.(6) with $Inv[h]$, and use Eq.(7) to obtain:

$$Inv[h] * y = Inv[h] * h * u = \delta * u = u \tag{9}$$

With the input u thus determined, the inverse problem for a linear time-invariant system can always be solved, provided the convolution inverse $Inv[h]$ to the system's impulse response can be evaluated. Unfortunately, in a number of cases, the convolution inverse cannot be determined within the space of smooth point-functions, much in the same way natural numbers do not have natural number inverses to the operation of addition: one either has to construct a new operation (ie. subtraction) or extend the space of natural numbers to the space of positive *and negative* integers:

$$n - n = 0, \quad \text{or} \tag{10}$$

$$n + (-n) = 0 \tag{11}$$

By analogy, the construction of a consistent convolution inverse to any impulse response will require:

- either a new defining operation, replacing convolution as the operation of choice for modelling LTI systems,
- or the extension of the space of smooth point-functions to a larger space, which will include more singular functions than just smooth or piecewise smooth point-functions.

This picture will become clearer in the next section, when Fourier transforms are introduced, and where it will be shown that there exists no smooth point-function such that convolving with any Gaussian $e^{-x^2/\lambda^2}/\sqrt{\pi}\lambda$ yields a Dirac delta; in other words, that there exists no smooth or piecewise smooth point-function that is convolution inverse to a Gaussian (of infinite support, i.e. defined in $]-\infty, +\infty[$).

1.2 Fourier Transform Solutions to the Inverse Problem

Fourier transforms are frequently utilized to solve inverse problems for linear time-invariant systems. That is because the operation of convolution gets mapped into a simple algebraic multiplication under the Fourier transform. If $u(x)$, $y(x)$, and $h(x)$ are respectively the input, output, and impulse response of a linear time-invariant system, and if $\tilde{y}(k)$ denotes the Fourier transform of $y(x)$:

$$\tilde{y}(k) = \int_{-\infty}^{+\infty} y(x)e^{-ikx}\,dx \tag{12}$$

$$y(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{y}(k)e^{+ikx}\,dk \tag{13}$$

then

$$y(x) = u(x) * h(x) \tag{14}$$

$$\tilde{y}(k) = \tilde{u}(k)\cdot\tilde{h}(k) \tag{15}$$

and we can thus obtain an expression for the input, given the output and the impulse response:

$$\tilde{u}(k) = \frac{\tilde{y}(k)}{\tilde{h}(k)} \tag{16}$$

$$u(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{\tilde{y}(k)}{\tilde{h}(k)}e^{+ikx}\,dk \tag{17}$$

However, eventual zeros of $\tilde{h}(k)$, at specific points in wave-number space, or at the limits $\pm\infty$ of the Fourier integral, will ill-condition the inverse-Fourier integral in Eq.(17). This result can be interpreted in the following way. In Fourier space, the blurred waveform is the product of the *optical transfer function* (OTF, Fourier transform of the point-spread function) with the original waveform. As the support of the point-spread function increases, the support of the optical transfer function decreases (a property of Fourier transforms), thus cutting off an increasing amount of high frequencies of the original waveform, which unavoidably "smooths" or "blurs" it. This process is depicted in figure 2. The inverse Fourier integral in Eq.(17) has to restore the high frequencies lost in the blurring process; the higher the frequencies, the greater the magnitude of the restoration. This process eventually leads to a diverging inverse-Fourier integral.

We now consider the case of a Gaussian filter $G(x,t)$:

$$G(x,t) = \frac{e^{-\frac{x^2}{4t}}}{\sqrt{4\pi t}} \tag{18}$$

and attempt to evaluate an explicit representation of the convolution inverse of $G(x,t)$ with the help of Fourier transforms. We suppress the spatial variable x if confusion does not arise, and denote $G(x,t)$ by $G(t)$. Thus, using standard definitions[2][3], we introduce

$$F(a,x,2t) = \frac{1}{2\pi}\int_{-a}^{+a} e^{-ikx+k^2t}\,dk \tag{19}$$

$$= \frac{e^{\frac{x^2}{4t}}}{4i\sqrt{\pi t}}[\mathrm{erf}(i\sqrt{t}\,a + \frac{x}{2\sqrt{t}}) + \mathrm{erf}(i\sqrt{t}\,a - \frac{x}{2\sqrt{t}})] \tag{20}$$

A graph of $F(a,x,1)$ appears in figure 3, and the time history of $F(a,x,t)$ appears in figure 4. In constructing the figures, the Nyquist criterion was applied to ensure that the factor e^{k^2t} did not generate frequency aliasing[1].

Proceeding formally, $Inv[G(t)] = \lim_{a\to\infty} F(a,x,2t)$, and it can be seen from figures 3 and 4 that the expression for $Inv[G(t)]$ diverges as the upper and lower integral limits $\pm a$ are increased in wave-number space. Nevertheless, engineering approximations of $Inv[G(t)]$ are frequently used by cutoff of the limits of the Fourier integral (also called finite-bandwidth approximations, since the support of the gaussian filter is reduced from $]-\infty, +\infty[$ to [-a,+a]). In other words, information on the impulse response and the output for large wave-numbers k has to be discarded. However, the approximations to the original input thus obtained are highly sensitive to the cutoff value a, and if there is no a priori knowledge of the Fourier spectrum width of the input, the complete recovery of the original input through Fourier transforms is an ill-posed problem. Nevertheless, Gaussians are impulse responses to a wide variety of physical systems[1], generally called *blurring systems*. Systems with pure Gaussian impulse response are responsible for gaussian blur, while systems without a Gaussian impulse response are responsible for non-gaussian blur. Telescopes and a vast majority of optical systems are blurring systems in two spatial dimensions. Communication satellites are blurring systems in one dimension (waveforms are transmitted sequentially). If the convolution inverse of a Gaussian (of infinite support) cannot be defined as a smooth point-function, how can it be defined and subsequently utilized to solve the inverse problem?

We propose here a novel approach to Gaussian deconvolution by constructing a rigorous relation between the deconvolution of a waveform that results from a Gaussian filter, and the integration of the diffusion equation with negative diffusivity (antidiffusion). In constructing this relationship, a new class of highly singular functions that

[1] or at least good approximations thereof

are called *hyperdistributions* are introduced, as well as their algebraic properties and a sequence of analyzing smooth point-functions (Hermite-Rodriguez wavelets) that approximate them.

2 Singular Functions: Distributions and Hyperdistributions

2.1 Taylor and Moment Expansions

The theory of hyperdistributions is built on ideas related to Dirac's delta function, which is technically a *distribution*[4]-[11]. The Dirac delta function, $\delta(x)$, has the following properties:

$$\int_{-\infty}^{+\infty} \delta(x)\, dx = 1 \tag{21}$$

and Dirac's identity:

$$f(x) = \int_{-\infty}^{+\infty} \delta(x - x') f(x') dx' \tag{22}$$

for any suitably smooth function $f(x)$.

Distributions, like the Dirac delta function, were first introduced by L. Schwartz[4], and widely publicized by G. Temple[6] and M. Lighthill[10]. In fact, Lighthill dedicated his book to Dirac, Schwartz, and Temple, in the following way:

To
PAUL DIRAC
who saw that it must be true,
LAURENT SCHWARTZ
who proved it,
AND
GEORGE TEMPLE
who showed how simple it could be made

Both Dirac and Temple are Englishmen, while Schwartz is a Frenchman. The English do have a knack for putting down the French, do they not?

Distributions allow for considerable simplification and increased mathematical elegance in the handling of integral and differential equations. For example, point-forces, or short-time impulses, are frequently mathematically modelled as Dirac delta functions. All expressions involving distributions, are assumed to hold under integration with any *test function*[6][10]. More specifically, distributions are defined by a sequence of functions, and the property of "weak convergence". Good functions (total

functions that are differentiable to all orders, and taper at $\pm\infty$ faster than any power) play the role of "testing" a sequence of functions for weak convergence. That is, a sequence of good functions $\{f_n(x)\}_n$ is a distribution if

$$\lim_{n\to\infty}\int_{-\infty}^{+\infty}\phi(x)f_n(x)\,dx < \infty \quad \text{for all good functions } \phi \tag{23}$$

For a more detailed introduction of distributions, refer to section 2.4: * Proof of Consistency: Taylor Expansion of the Delta ' unction.

We now expand $\delta(x'-x)$ in a Taylor expansion[2], in terms of its derivatives about x':

$$\delta(x'-x) = \delta(x') - x\delta'(x') + \frac{x^2}{2!}\delta''(x') + \ldots \tag{24}$$

where $\delta^n(x) = \frac{d^n}{dx^n}\delta(x)$. A "local" approximation to $f(x)$ can be derived from Eq.(24) in the following way. By substituting Eq.(24) into Dirac 's identity, we recover the usual Taylor expansion of $f(x)$ about $x = 0$:

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \ldots + R^n(x) \tag{25}$$

This approximation is local in the sense that it requires derivatives of $f(x)$ at a single point, $x = 0$, and in general has a limited radius of convergence. On the other hand, by expanding $\delta(x-x')$ in a Taylor expansion about x:

$$\delta(x-x') = \delta(x) - x'\delta'(x) + \frac{x'^2}{2!}\delta''(x) + \ldots \tag{26}$$

When this series is substituted into Dirac 's identity, We obtain

$$f(x) = M^0\delta(x) - M^1\delta'(x) + \frac{M^2}{2!}\delta''(x) + \ldots + \hat{R}^n(x) \tag{27}$$

The coefficients M^n, defined by

$$M^n = \int_{-\infty}^{+\infty} x^n f(x)\,dx \tag{28}$$

are the centered moments of the function $f(x)$. Therefore, Eq.(27) is an approximation of $f(x)$ involving global information about $f(x)$.

The global expansion of $f(x)$ in terms of its centered moments and derivatives of the Dirac delta function is the motivation for the introduction of hyperdistributions.

[2]this is accomplished in a formal way; the justification and proof-of-consistency will be given in section 2.4: * Proof of Consistency: Taylor Expansion of the Delta Function

Definition (formal): Any "function" which may be formally written as,

$$f(x) = \sum_{n=0}^{\infty} a_n \delta^n(x) \tag{29}$$

with finite and real coefficients a_n given by

$$a_n = (-1)^n \frac{M^n}{n!} \quad \text{with} \quad M^n = \int_{-\infty}^{+\infty} x^n f(x)\, dx \tag{30}$$

defines a hyperdistribution.

To carry on the analogy with natural numbers introduced previously, hyperdistributions can be identified with the set of negative integers. Much in the same way hyperdistributions are not *smooth* point-functions, negative integers do not correspond to any *physical* reality. Indeed, we can physically represent the integer 2 by picturing, say, two apples (or physical entities of any other sort) in our mind. But how do we represent the negative integer -2? Is an "absence" a constructive definition? In spite of this conceptual difficulty, negative integers are required in simple arithmetic, since they are inverses to the set of natural numbers for the operation of addition. That is the reason why the space of natural numbers is extended to the space of positive *and* negative integers. In a similar fashion, hyperdistributions do not correspond to any physical reality (e.g. hyperdistributions are not smooth –or even piecewise smooth– point-functions), yet it will be shown that hyperdistributions are convolution inverses to smooth point-functions, and are thus required in the "convolution calculus" of signals and systems. One is thus naturally led to extend the space of smooth point-functions to include highly singular functions such as hyperdistributions.

2.2 The Convolution Group

Given any two hyperdistributions f_1 and f_2, their linear combination $\lambda f_1 + \mu f_2$ is also a hyperdistribution (where λ and μ are real numbers):

$$f_1(x) = \sum_{n=0}^{\infty} a_n \delta^n(x) \tag{31}$$

$$f_2(x) = \sum_{n=0}^{\infty} b_n \delta^n(x) \tag{32}$$

$$\lambda f_1(x) + \mu f_2(x) = \sum_{n=0}^{\infty} (\lambda a_n + \mu b_n) \delta^n(x) \tag{33}$$

The p^{th} derivative $d^p f(x)/dx^p = \nabla^p f(x)$ of a hyperdistribution $f(x)$ is also a hyperdistribution:

$$\nabla^p f(x) = \nabla^p \sum_{n=0}^{\infty} a_n \delta^n(x) = \sum_{n=0}^{\infty} b_n \delta^n(x) \quad \text{where} \quad b_n = \begin{cases} 0 & \text{if } n < p \\ a_{n-p} & \text{otherwise} \end{cases} \tag{34}$$

The convolution of two hyperdistributions $f_1 * f_2$ is a hyperdistribution:

$$f_1(x) * f_2(x) = \sum_{n=0}^{\infty} a_n \nabla^n \delta(x) * \sum_{n=0}^{\infty} b_n \nabla^n \delta(x) \tag{35}$$

$$= \int_{-\infty}^{+\infty} (\sum_{p=0}^{\infty} a_p \nabla_{x'}^p \delta(x'))(\sum_{q=0}^{\infty} b_q \nabla_{x-x'}^q \delta(x-x'))\, dx' \tag{36}$$

Noting that $\nabla_{x-x'}^q = (-1)^q \nabla_{x'}^q$, and with q integrations by parts, we obtain:

$$f_1(x) * f_2(x) = (-1)^{2q} \sum_{p=0}^{\infty} (a_p \sum_{p=0}^{\infty} b_q \int_{-\infty}^{+\infty} \nabla_{x'}^{p+q} \delta(x') \delta(x-x')\, dx') \tag{37}$$

From Dirac's identity,

$$\int_{-\infty}^{+\infty} \nabla_{x'}^{p+q} \delta(x') \delta(x-x')\, dx' = \nabla^{p+q} \delta(x) \tag{38}$$

and thus Eq.(36) may finally be reduced to

$$f_1(x) * f_2(x) = \sum_{p=0}^{\infty} (a_p \sum_{q=0}^{\infty} b_q \nabla^{p+q} \delta(x)) \tag{39}$$

or, equivalently, with $r = p + q$:

$$f_1(x) * f_2(x) = \sum_{r=0}^{\infty} \left\{ \sum_{p=0}^{r} a_p b_{r-p} \right\} \nabla^r \delta(x) \tag{40}$$

Since $\{a_n\}_n$ and $\{b_n\}_n$ are real and finite, $\left\{\sum_{p=0}^{r} a_p b_{r-p}\right\}_r$ are real and finite, and Eq.(40) defines a hyperdistribution. The three properties in Eqs.(33), (34), and (40) above may be thought of as "closure properties" of hyperdistributions.

Another attractive property of hyperdistributions is their behavior under the Fourier transform. Again taking $f(x) = \sum_{n=0}^{\infty} a_n \nabla^n \delta(x)$, the Fourier transform can be evaluated as:

$$\int_{-\infty}^{+\infty} e^{-ikx} f(x)\, dx = \sum_{n=0}^{\infty} a_n \int_{-\infty}^{+\infty} e^{-ikx} \nabla^n \delta(x)\, dx \tag{41}$$

$$= \sum_{n=0}^{\infty} a_n (-1)^n \int_{-\infty}^{+\infty} (\nabla^n e^{-ikx}) \delta(x)\, dx \tag{42}$$

$$= \sum_{n=0}^{\infty} a_n (ik)^n \tag{43}$$

where we have used integration by parts n times, and Dirac's identity with $\nabla^n e^{ikx} \mid_{x=0} = (ik)^n$. The Fourier transform of a hyperdistribution is thus a formal power series in

the wave-number k. Moreover, the Fourier transform of a function is the "moment generating function", because:

$$a_n = (-1)^n \frac{M^{(n)}}{n!} \quad \text{with} \quad M^{(n)} = \int_{-\infty}^{\infty} x^n f(x)\, dx \tag{44}$$

As a result, the definition of hyperdistributions, which requires that the coefficients a_n be real and finite, is equivalent to requiring that its Fourier transform be a real analytic function of the wave-number k.

Hyperdistributions allow an effective computation of the convolution inverse. Given a hyperdistribution f, the desired convolution inverse $Inv[f]$ satisfies

$$f * Inv[f] = \delta \tag{45}$$

where δ represents Dirac 's delta function, which, by Dirac's identity, is the unit element of the convolution operation. We shall show by construction that if f can be written as a hyperdistribution, then $Inv[f]$ is also a hyperdistribution. By definition,

$$(f * Inv[f])(x) = \int_{-\infty}^{\infty} f(x') Inv[f](x - x')\, dx' = \delta(x) \tag{46}$$

It is now necessary to compute the product of the sums and match coefficients. Taking $f(x) = \sum_n a_n \nabla^n \delta(x)$, and $Inv[f(x)] = \sum_n b_n \nabla^n \delta(x)$, it is seen that the computation of the convolution inverse is effectively the determination of a collection of b_n values, given a set of a_n values. Substituting into Eq.(46),

$$\int_{-\infty}^{+\infty} (\sum_{p=0}^{\infty} a_p \nabla'^p \delta(x'))(\sum_{q=0}^{\infty} b_q \nabla^q \delta(x - x'))\, dx' = \delta(x) \tag{47}$$

Once again noting that $\nabla^q_{x-x'} = (-1)^q \nabla_{x'}$ and that $\nabla^p \delta(x) * \nabla^q \delta(x) = \nabla^{p+q} \delta(x)$, Eq.(47) may finally be reduced to:

$$\sum_{r=0}^{\infty} \left\{ \sum_{p=0}^{r} a_p b_{r-p} \right\} \nabla^r \delta(x) = \delta(x) \tag{48}$$

Matching coefficients on the left-hand and right-hand sides implies that only the $r = 0$ term survives. The result is a linear system of equations for the b values in terms of the a values. It is easiest to see the behavior by writing the first few equations in this linear system,

$$\begin{aligned} a_0 b_0 &= 1 \\ a_0 b_1 + a_1 b_0 &= 0 \\ a_0 b_2 + a_1 b_1 + a_2 b_0 &= 0 \\ a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0 &= 0 \end{aligned} \tag{49}$$

and so forth. Thus, it can be seen that, $b_0 = 1/a_0$, $b_1 = -a_1/a_0^2$ $b_2 = a_1^2/a_0^3 - a_2/a_0^2$, etc.. The linear system in Eq.(49) is called the *Bochner* algebra for hyperdistributions, after a very similar algebra, developed by S. Bochner[12] for power series.

In matrix notation, and if $a_0 = 1$, the coefficients b_n are given by the following Toeplitz[13] determinant:

$$b_n = (-1)^n \begin{vmatrix} a_1 & 1 & 0 & 0 & \dots & 0 \\ a_2 & a_1 & 1 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & 1 & \dots & 0 \\ & & & & & \\ \vdots & & & & \ddots & \vdots \\ & & & & & 1 \\ a_n & a_{n-1} & a_{n-2} & a_{n-3} & \dots & a_1 \end{vmatrix} \tag{50}$$

We note that $Inv[f]$ is indeed a hyperdistribution, and is determined by this algorithm up to a function with vanishing moments.

Some convolution inverses appear in the familiar context of potential theory. For example, the "isotropic quadrupole", which is a distribution, is the convolution inverse of the Coulomb potential in 3 dimensions:

$$\nabla^2 \delta(\vec{r}) = Inv[\frac{-1}{4\pi r}]$$

Other simple examples of convolution inverse pairs in one dimensional potential theory are $(\mid x \mid, \delta''(x))$, $(sgn(x), \delta'(x))$, and $(\delta(x), \delta(x))$. More examples of convolution inverses can be obtained from standard Green's functions. For example, the one-dimensional Helmholtz equation

$$\frac{d^2}{dx^2}u + p^2 u = 0 \tag{51}$$

can be analyzed with hyperdistributions, as well as with Fourier transforms, to obtain both its Green's function and the convolution inverse of its Green's function. Thus,

$$(\frac{d^2}{dx^2} + \frac{1}{4})e^{-|x|/2} = -\delta(x) \tag{52}$$

The Bochner algebra for hyperdistributions shows that only the terms b_0 and b_2 contribute to the convolution inverse of the Green's function, thus yielding:

$$Inv[e^{-|x|/2}] = \frac{1}{2}\delta(x) - \delta''(x) \tag{53}$$

The result above is easily reproduced by utilizing Fourier transforms. However, it was shown in Eq.(19) that the convolution inverse of a Gaussian cannot be obtained with

Fourier transforms. In contrast, that same convolution inverse can easily be obtained through the Bochner algebra for hyperdistributions: calculation of the power moments of a Gaussian $\delta_\lambda(x)$ of width λ

$$\delta_\lambda(x) = \frac{e^{-\frac{x^2}{\lambda^2}}}{\sqrt{\pi}\lambda} \tag{54}$$

yields:

$$\begin{aligned} a_{2n} &= \int_{-\infty}^{+\infty} x^{2n} e^{-\frac{x^2}{\lambda^2}} / \sqrt{\pi}\lambda \, dx \\ &= \lambda^{2n} \frac{(2n-1)!!}{2^n} \qquad (55) \\ a_{2n+1} &= 0 \qquad (56) \end{aligned}$$

where the double factorial involves exclusively the product of odd numbers. This leads to the hyperdistribution expansion of $\delta_\lambda(x)$:

$$\delta_\lambda(x) = \sum_{n=0}^{\infty} \left(\frac{\lambda}{\sqrt{2}}\right)^{2n} \frac{1}{(2n)!} \nabla^{2n} \delta(x) \tag{57}$$

The Bochner algebra for hyperdistributions yields the convolution inverse of the Gaussian $\delta_\lambda(x)$:

$$Inv[\delta_\lambda(x)] = \sum_{n=0}^{\infty} \left(-\frac{\lambda}{\sqrt{2}}\right)^{2n} \frac{1}{(2n)!} \nabla^{2n} \delta(x) \tag{58}$$

In a more compact operator notation:

$$\begin{aligned} \delta_\lambda(x) &= e^{+\lambda/\sqrt{2}\nabla^2} \delta(x) \qquad (59) \\ Inv[\delta_\lambda(x)] &= e^{-\lambda/\sqrt{2}\nabla^2} \delta(x) \qquad (60) \end{aligned}$$

In this section, it was shown that the convolution of any two hyperdistributions yields another hyperdistribution. Furthermore, it was proven that any hyperdistribution has another hyperdistribution as its convolution inverse, provided the impulse response can also be written as a hyperdistribution. A general algorithm for deconvolving any hyperdistribution –the Bochner algebra– was given. As a result, hyperdistributions provide a closure for the classical semi-group of smooth point-functions with the operation of convolution[14]. If $\mathcal{H}$ denotes the space of Hyperdistributions, with the space of smooth point functions embedded in it, $(\mathcal{H}, *)$ is an abelian group, with the Dirac delta as the unit element. Hyperdistributions are thus an extremely useful tool for solving integral equations of convolution-type[15].

2.3 Function-Theoretic Definition of Hyperdistributions

The process for defining hyperdistributions parallels the Temple definition[3] of a distribution (generalized function) as a good sequence of good functions[4]-[11]. Good functions are smooth and tapered. More precisely, they are total point functions that are differentiable to all orders (C^∞), and decay at $\pm\infty$ faster than any power. Good functions play the role of "testing" a sequence of good functions for weak convergence. In fact, a sequence of good functions is a distribution if

$$\lim_{n\to\infty}\int_{-\infty}^{+\infty}\phi(x)f_n(x)\,dx < \infty \quad \text{for all good functions } \phi \tag{61}$$

Since hyperdistributions are conceived as "generalized" distributions, a second order generalization of functions is in fact implemented. Consequently, a double test is needed as a convergence criterion. This criterion is implemented by introducing *very good* functions $G_\Lambda(x)$ with the following properties:

1. $G_\Lambda(x)$ is smooth, that is, differentiable to all orders, i.e. C^∞.
2. $G_\Lambda(x)$ is essentially compact, i.e. it has a Gaussian decay at $\pm\infty$: $G_\Lambda(x) \sim \lim_{x\to\infty} Ne^{-x^2/\Lambda^2}$.

We will assume for convenience that G_Λ is normalized to unity:

$$\int_{-\infty}^{+\infty} G_\Lambda(x)\,dx = 1 \tag{62}$$

The width of G_Λ is defined by

$$\frac{\Lambda^2}{4} = \int_{-\infty}^{+\infty}(x-\bar{x})^2 G_\Lambda(x)\,dx \tag{63}$$

A primary example of very good functions is the Gaussian, which is denoted by $\delta_\lambda(x)$:

$$\delta_\lambda(x) = \frac{e^{-\frac{x^2}{\lambda^2}}}{\sqrt{\pi}\lambda} \tag{64}$$

A sequence of very good functions is now introduced, defined by

$$\mathcal{H}_n^\lambda(x) = \sum_{k=0}^{n} a_k \nabla^k \delta_\lambda(x) \tag{65}$$

[3]Lighthill[10] points out that in so doing, Temple follows Mikusinski's[16][17] approach

The sequence $\{\mathcal{H}_n^\lambda\}_{\lambda,n}$, where λ is a nonnegative real and n is a natural number, is a good sequence if, for all good functions ϕ and for all very good functions G_Λ, there exists a Λ_0 such that, for all $\Lambda > \Lambda_0$,

$$\lim_{\substack{n\to\infty \\ \lambda\to 0}} \int_{-\infty}^{+\infty} \phi(x)\,(\mathcal{H}_n^\lambda * G_\Lambda)(x)\,dx < \infty \tag{66}$$

We note that $e^{\pm t\nabla^2}\delta(x)$ are hyperdistributions. The sum (hyperdistribution)

$$\sum_{k=0}^{\infty} a_k \nabla^k \delta(x) \tag{67}$$

can thus be viewed either as a sequence of "good" distributions as $n \to \infty$

$$\sum_{k=0}^{n} a_k \nabla^k \delta(x) \tag{68}$$

or, as $\lambda \to 0$, as a sequence of good functions:

$$\sum_{k=0}^{\infty} a_k \nabla^k \delta_\lambda(x) \tag{69}$$

The latter representation is a *Rodriguez expansion.* The Rodriguez formula for Hermite polynomials[2] can be used to show that the derivatives of a Gaussian form a complete set of orthogonal polynomials in an L^2 space. And thus the Rodriguez expansion yields a very useful point function approximation to any hyperdistribution:

$$\sum_{k=0}^{\infty} a_k\,(-1)^n \frac{H_n(\frac{x}{\lambda})}{\lambda^n}\,\delta_\lambda(x) \tag{70}$$

where $H_n(x)$ denotes the Hermite Polynomial in x of order n.

The problem of "approximating" hyperdistributions with smooth point-functions will be studied shortly.

2.4 * Proof of Consistency: Taylor Expansion of the Delta Function

We start by quoting three useful definitions utilized by Lighthill[10] in his definition of distributions[4].

[4] Lighthill actually uses the terminology "generalised function", instead of "distributions", which was first introduced by Schwartz[4]

Definition 1 (p. 15). A good function is one which is everywhere differentiable any number of times and such that it and all its derivatives are $\mathcal{O}(\mid x \mid^{-N})$ as $\mid x \mid \rightarrow \infty$ for all N.

Example 1. e^{-x^2} is a good function.

Definition 2. A fairly good function is one which is everywhere differentiable any number of times and such that it and all its derivatives are $\mathcal{O}(\mid x \mid^{-N})$ as $\mid x \mid \rightarrow \infty$ for some N.

Example 2. Any polynomial is a fairly good function.

Definition 3. A sequence $f_n(x)$ of good functions is called regular if, for any good function $F(x)$ whatever, the limit

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} f_n(x)F(x)dx \tag{71}$$

exists.

(...) Definition 5. A generalised function $f(x)$ is defined as a regular sequence $f_n(x)$ of good functions (...). The integral

$$\int_{-\infty}^{+\infty} f(x)F(x)dx \tag{72}$$

of the product of a generalised function $f(x)$ and a good function $F(x)$ is defined as

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} f_n(x)F(x)dx \tag{73}$$

In other words, Lighthill defines distributions by a sequence of good functions (total functions, differentiable to all orders, that taper at $\pm\infty$ faster than any power), and by the property of weak convergence, which requires that the limit in Eq.(71) exists, *for any* choice of good (testing) function $F(x)$.

We now justify the expansion in Eq.(24) of $\delta(x' - x)$ in a Taylor series, in terms of its derivatives about x' (since $\delta(x' - x)$ is technically a distribution, and not a point-function, such a justification and proof-of-consistency is required). The Taylor expansion-with-remainder of $\delta(x' - x)$ about x':

$$\delta(x' - x) = \delta(x') + x\delta'(x') + \frac{x^2}{2!}\delta''(x') + \ldots + \frac{x^{N-1}}{(N-1)!}\delta^{(N-1)}(x') + R^N \tag{74}$$

To be consistent with Lighthill's definition of distributions, We are required to test Eq.(74) for <u>all</u> available good functions $\phi(x)$:

$$\int_{-\infty}^{+\infty} \delta(x' - x)\phi(x')dx' = \int_{-\infty}^{+\infty} \left\{ \sum_{n=0}^{N-1} \frac{x^n}{n!}\delta^{(n)}(x') + R^N \right\} \phi(x')dx' \tag{75}$$

By Dirac's identity (Eq.(22)),

$$\phi(x) = \sum_{n=0}^{N-1} \frac{x^n}{n!}\phi^{(n)}(0) + \mathcal{R}^N \quad \text{where} \quad \mathcal{R}^N = \int_{-\infty}^{+\infty} R^N \phi(x')dx' \tag{76}$$

and we recognize Taylor's expansion for the test function $\phi(x)$. If such an expansion is indeed legitimate for all test functions $\phi(x)$, then Eq.(74) is justified and consistent. However, Lighthill does not require of test functions that their Taylor expansions converge, or, in other words, that all test functions be real-analytic.

> *Theorem of Consistency for Hyperdistributions*: In restricting the space of test functions to good functions that are real-analytic (i.e. C^ω), Eq.(74) is justified, and the introduction of the hyperdistribution sum in Eq.(29) is consistent.

It might seem paradoxical that by restricting the space of Lighthill test function (from C^∞ to C^ω), we expand Lighthill's space of singular functions to include hyperdistributions, which are even more singular than distributions. In fact, this is a consequence of the "dual" behavior of Lighthill's regular sequences and Lighthill's test functions under the weak convergence property. Since the limit

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} f_n(x)F(x)dx \tag{77}$$

must exist, extending the space of regular sequences $f_n(x)$ necessarily restricts the space of available test functions $F(x)$.

2.5 Approximating Hyperdistributions with Smooth Point-Functions

In this section, we will parallel Lighthill's approach in defining the delta function by a sequence of narrowing gaussians. We will introduce a sequence of smooth point-functions, whose limit is in fact a hyperdistribution. These smooth point-functions will then be the focus of section 3: Hermite-Rodriguez Wavelet Analysis.

2.5.1 "Diffusing" Hyperdistributions

It is observed that by formally convolving a hyperdistribution $\sum_{n=0}^{\infty} a_n \nabla^n \delta(x)$ with a Gaussian of width λ:

$$\delta_\lambda(x) = \frac{e^{-x^2/\lambda^2}}{\sqrt{\pi}\lambda} \tag{78}$$

one obtains an infinite series of derivatives of Gaussians:

$$\sum_{n=0}^{\infty} a_n \nabla^n \delta(x) \ * \ \delta_\lambda(x) = \sum_{n=0}^{\infty} a_n \{\nabla^n \delta(x) \ * \ \delta_\lambda(x)\} = \sum_{n=0}^{\infty} a_n \nabla^n \delta_\lambda(x) \qquad (79)$$

The first equality is a formal one. The proof of the second equality follows from the application of the chain rule to the operator $\nabla^n_{x-x'}$, n integration by parts, and Dirac's identity:

$$\begin{aligned} \nabla^n \delta(x) * \delta_\lambda(x) &= \int_{-\infty}^{+\infty} \nabla^n_{x-x'} \delta(x-x') \delta_\lambda(x')\, dx' \\ &= (-1)^n \int_{-\infty}^{+\infty} \nabla^n_{x'} \delta(x-x') \delta_\lambda(x')\, dx' \\ &= (-1)^{2n} \int_{-\infty}^{+\infty} \delta(x-x') \nabla^n_{x'} \delta_\lambda(x')\, dx' \\ &= \nabla^n \delta_\lambda(x) \end{aligned} \qquad (80)$$

In other words, diffusion "maps" hyperdistributions, which are singular functions, into series of smooth point-functions. This result is not surprising. Diffusion is a smoothing operation, and is expected to map singular functions into smooth functions. This is indeed the case for Schwartz distributions: Picture a single Dirac delta as the initial condition (at $t = 0$) of a diffusion process in time. The 1+1 homogeneous diffusion equation can be written as

$$\frac{\partial u}{\partial t} - \nabla^2 u = 0 \qquad (81)$$

The initial-condition Green's function $G_t(x)$ for such a diffusion equation is the Gaussian[22] $G_t(x) = \delta_{2\sqrt{t}}(x)$. In other words, if $f(x)$ is the initial condition to the diffusion process, the solution of the diffusion equation at time t can be written as:

$$u(x,t) = G_t(x) \ * \ f(x) = \delta_{2\sqrt{t}} \ * \ f(x) \qquad (82)$$

or, in operator form[23],

$$u(x,t) = e^{t\nabla^2} f(x) \qquad (83)$$

Since in this case $f(x) = \delta(x)$, we have

$$u(x,t) = G_t(x) \ * \ \delta(x) = \delta_{2\sqrt{t}}(x) \ * \ \delta(x) = \delta_{2\sqrt{t}}(x) \qquad (84)$$

The last equality follows once again from Dirac's identity. In other words, at $t = \epsilon$, for any $\epsilon > 0$ however small, a Dirac delta is "mapped" into a Gaussian of width $\sqrt{2t}$ by the diffusion equation. Similarly, it can be proven using Eq.(80) that a finite derivative of a Dirac delta, gets mapped to the derivative of same order of a Gaussian of width $\sqrt{2t}$, by the diffusion equation:

$$u(x,t) = G_t(x) \ * \ \nabla^p \delta(x) = \delta_{2\sqrt{t}}(x) \ * \ \nabla^p \delta(x) = \nabla^p \delta_{2\sqrt{t}}(x) \qquad (85)$$

We now return to hyperdistributions, by picturing an infinite number of Dirac delta functions, all centered at the origin, and with different amplitudes a_n, that is: $\sum_{n=0}^{\infty} a_n \nabla^n \delta(x)$, as an initial condition to a diffusion process. That *hyperdistribution* will formally *diffuse* at time $t > 0$, however small t may be, into the following functional series:

$$\sum_{n=0}^{\infty} a_n \nabla^n \delta_{2\sqrt{t}}(x) \tag{86}$$

In other words, diffusion maps hyperdistributions into series of smooth point-functions. If the series converge, we can say that diffusion maps hyperdistributions onto the space of smooth point-functions. The sufficient conditions for convergence of the series in Eq.(86) remains however to be investigated.

To this end, the Rodriguez formula for Hermite Polynomials[2]:

$$(-1)^n H_n(x) e^{-x^2} = \nabla^n e^{-x^2} \tag{87}$$

can be rescaled as follows:

$$(-1)^n H_n^{\lambda}(x) \delta_{\lambda}(x) = \nabla^n \delta_{\lambda}(x) \tag{88}$$

where

$$H_n^{\lambda}(x) = \frac{H_n(\frac{x}{\lambda})}{\lambda^n} \tag{89}$$

And thus Eq.(86) becomes a series involving Hermite Polynomials:

$$\sum_{n=0}^{\infty} a_n (-1)^n H_n^{2\sqrt{t}}(x) \delta_{2\sqrt{t}}(x) \tag{90}$$

Unfortunately, the series above is *not* a classical Hermite series, since the coefficients a_n are centered power-moments given by Eq.(30) (weighted by monomials in x, and *not* by Hermite polynomials). As a result, one cannot use the standard tools of Christoffel-Darboux theory[2], which yield the sufficient conditions for the convergence of Hermite series[2]. The task that lies ahead is to transform the series in Eq.(90) into standard Hermite series by relating the centered power-moments $\int_{-\infty}^{+\infty} x^n f(x)\, dx$ to the centered Hermite moments $\int_{-\infty}^{+\infty} H_n(x) f(x)\, dx$, where $H_n(x)$ denotes the Hermite polynomial of order n. That is the purpose of the C-matrix transform, which will be studied shortly.

2.5.2 Hermite-Rodriguez Expansions

We now expand the function $f(x)$ directly in terms of Hermite polynomials. Hermite polynomials constitute an orthogonal set of basis functions for the appropriate L^2 space, and such an expansion is thus legitimate. We start with the expansion of $f(x)$

in terms of derivatives of the Gaussian of width λ, with yet unknown coefficients (or moments) b_n:

$$f(x) = \sum_{n=0}^{+\infty} b_n \nabla^n \delta_\lambda(x) \tag{91}$$

where

$$\delta_\lambda(x) = \frac{e^{-\frac{x^2}{\lambda^2}}}{\sqrt{\pi}\lambda} \tag{92}$$

With the use of the Rodriguez formula for scaled Hermite polynomials derived in Eq.(88):

$$(-1)^n H_n^\lambda(x)\delta_\lambda(x) = \nabla^n \delta_\lambda(x) \tag{93}$$

with

$$H_n^\lambda(x) = \frac{H_n(x/\lambda)}{\lambda^n} \tag{94}$$

We substitute$\nabla^n \delta_\lambda(x)$ in Eq.(91) with the expression in Eq.(93). Eq.(91) becomes:

$$f(x) = \sum_{n=0}^{+\infty} b_n (-1)^n H_n^\lambda(x)\delta_\lambda(x) \tag{95}$$

We now utilize the orthogonality of Hermite polynomials:

$$\int_{-\infty}^{+\infty} H_n(x)H_m(x)e^{-x^2}dx = \sqrt{\pi}2^n n!\delta_{nm} \tag{96}$$

where δ_{nm} represents the Kronecker delta, which is 0 if $n \neq m$ and 1 if $n = m$. We multiply Eq.(95) with $H_m^\lambda(x)$ and integrate from $-\infty$ to $+\infty$. Assuming the sum in Eq.(95) converges, We can interchange the integral and sum symbols, and utilize Eq.(96) to obtain:

$$(-1)^n b_n = \frac{\lambda^{2n}}{2^n n!}\int_{-\infty}^{+\infty} f(x)H_n^\lambda(x)dx \tag{97}$$

The coefficients b_n above are called the *Hermite-Rodriguez moments* of the Hermite-Rodriguez expansion of $f(x)$:

$$f(x) = \sum_{n=0}^{\infty}(-1)^n b_n H_n^\lambda(x)\delta_\lambda(x) \tag{98}$$

hereafter denoted as the HR expansion of $f(x)$ with weight λ. The weight parameter is a novel feature of HR expansions, when compared to standard expansions in terms of complete sets of basis functions. The standard expansions do not contain a free parameter. This parameter may be employed to speed-up the convergence of HR expansions. In other words, for every function $f(x)$ to be approximated by an HR expansion, there exists one (or possibly more) choice(s) of the weight λ for which the HR expansion converges to a good approximation of $f(x)$ with a minimum number of partial sums.

2.5.3 Power Moments, Hermite-Rodriguez Moments, and the $\mathcal{C}$-Matrix Transform

We now relate the centered power moments of a function $f(x)$:

$$a_n = \frac{(-1)^n}{n!} \int_{-\infty}^{+\infty} x^n f(x)\, dx \tag{99}$$

with the Hermite-Rodriguez moments of the same function:

$$b_n^\lambda = \frac{(-1)^n \lambda^{2n}}{2^n n!} \int_{-\infty}^{+\infty} H_n^\lambda(x) f(x)\, dx \tag{100}$$

By utilizing the classical power-series expansion of Hermite polynomials[2]:

$$H_n(x) = n! \sum_{p=0}^{[n/2]} (-1)^p \frac{2^{n-2p}}{p!(n-2p)!} x^{n-2p} \tag{101}$$

where $[n/2]$ denotes the natural number inferior or equal to the rational $n/2$, and by substituting Eq.(101) in Eq.(100), we obtain:

$$b_n^\lambda = \frac{(-1)^n \lambda^n}{2^n} \sum_{p=0}^{[n/2]} (-1)^p \frac{2^{n-2p}}{p!(n-2p)!} \int_{-\infty}^{+\infty} (x/\lambda)^{n-2p} f(x)\, dx \tag{102}$$

and thus,

$$b_n^\lambda = \sum_{p=0}^{[n/2]} (-1)^p \frac{\lambda^{2p}}{p! 2^{2p}} a_{n-2p} \tag{103}$$

In matrix notation:

$$\begin{pmatrix} b_0^\lambda \\ b_1^\lambda \\ b_2^\lambda \\ \vdots \end{pmatrix} = \mathcal{C}^\lambda . \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \end{pmatrix} \tag{104}$$

where the matrix $\mathcal{C}^\lambda$ is lower-triangular, and is given by:

$$\mathcal{C}_{ij}^\lambda = \begin{cases} (-1)^{\frac{i-j}{2}} \frac{\lambda^{i-j}}{2^{i-j}(\frac{i-j}{2})!} & \text{if } (i-j) \text{ is even and positive} \\ 0 & \text{otherwise} \end{cases} \tag{105}$$

The determinant of $\mathcal{C}^\lambda$ is always equal to 1. As a result, one can always transform an HR expansion of $f(x)$ into a power-moment series expansion of the same function, and vice-versa. The problem of convergence of the series in Eq.(90) has effectively been reduced to the problem of convergence of Hermite expansions. Christoffel-Darboux theory[2] then yields sufficient conditions for convergence of Hermite-Rodriguez expansions.

2.5.4 Approximating Hyperdistributions with Hermite-Rodriguez Expansions

$f(x)$ is expanded as a hyperdistribution:

$$f(x) = \sum_{n=0}^{\infty} a_n \nabla^n \delta(x) \tag{106}$$

where

$$a_n = \frac{(-1)^n}{n!} \int_{-\infty}^{+\infty} x^n f(x)\, dx \tag{107}$$

That hyperdistribution is then approximated with a Hermite-Rodriguez series of weight λ:

$$f(x) = \sum_{n=0}^{\infty} b_n^{\lambda} \nabla^n \delta_{\lambda}(x) \tag{108}$$

where the Hermite-Rodriguez moments b_n are given by the two equivalent equations below:

$$b_n^{\lambda} = \frac{(-1)^n \lambda^{2n}}{2^n n!} \int_{-\infty}^{+\infty} f(x) H_n^{\lambda}(x)\, dx \tag{109}$$

$$a_n^{\lambda} = \frac{(-1)^n}{n!} \int_{-\infty}^{+\infty} x^n f(x)\, dx$$

$$\text{with } b_n^{\lambda} = \left(C_{nm}^{\lambda}\right) a_m \tag{110}$$

The sufficient condition for convergence of Hermite-Rodriguez expansions is given in the following section.

2.5.5 Convergence of Hermite-Rodriguez Expansions

Christoffel-Darboux theory[2] yields the sufficient condition for convergence of Hermite series:

$$\sum_{n=0}^{\infty} c_n H_n(x) \quad \text{with} \quad c_n = \frac{1}{2^n n! \sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} f(x) H_n(x)\, dx \tag{111}$$

converges to $f(x)$ if the following condition is met[2]:

$$\int_{-\infty}^{+\infty} e^{-x^2} f^2(x)\, dx < \infty \tag{112}$$

Hermite-Rodriguez expansions are of the form

$$\sum_{n=0}^{\infty} b_n \frac{H_n(x/\lambda)}{\lambda^n} \delta_{\lambda}(x) \quad \text{with} \quad b_n = \frac{\lambda^{2n}}{2^n n!} \int_{-\infty}^{+\infty} f(x) \frac{H_n(x/\lambda)}{\lambda^n}\, dx \tag{113}$$

By multiplying Eq.(111) by $\frac{\delta_\lambda(x)}{\lambda^n}$ and by rescaling from x to x/λ, the convergence criterion for Hermite Rodriguez expansions is obtained. It is the following sufficient condition. The series in Eq.(113) converges if the following condition is met:

$$\int_{-\infty}^{+\infty} e^{x^2/\lambda^2} f^2(x)\, dx < \infty \tag{114}$$

2.6 Hyperdistributions in 2 Dimensions

The generalization of hyperdistributions to two spatial dimensions is now introduced. Formal expressions of the type

$$f(x,y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_{nm}\, \nabla_x^n \nabla_y^m\, \delta(x,y) \tag{115}$$

which are called hyperdistributions, form an algebraic field. ∇_x^n and ∇_y^m respectively denote the n^{th} and m^{th} partial derivative operators in the x and y directions, and $\delta(x,y)$ denotes the 2D Dirac delta function:

$$\delta(x,y) = \delta(x)\delta(y) \tag{116}$$

The coefficients a_{nm} are given by

$$a_{nm} = \frac{(-1)^{n+m}}{n!m!} \iint_{-\infty}^{+\infty} x^n y^m f(x,y)\, dxdy \tag{117}$$

If $f(x,y)$ is convolved with

$$g(x,y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} b_{nm}\, \nabla_x^n \nabla_y^m\, \delta(x,y) \tag{118}$$

through the usual convolution product $*$, the result is once again a hyperdistribution, with coefficients given by

$$c = a \star b \tag{119}$$

This is the discrete convolution product, which can be written explicitly as

$$c_{nm} = \sum_{p=0}^{n} \sum_{q=0}^{m} a_{n-p\,,\,m-q} b_{pq} \tag{120}$$

To prove the above formula, we observe that

$$(\nabla_x^n \nabla_y^m\, \delta(x,y)) * (\nabla_x^p \nabla_y^q\, \delta(x,y)) = \nabla_x^{n+p} \nabla_y^{m+q}\, \delta(x,y) \tag{121}$$

which can be verified by repeated integration by parts. The algebra represented by Equation (120) is the 2D *Bochner algebra*[12]. This constitutes an efficient means to

deconvolve 2D images. In particular, the coefficients of the convolution product are given by

$$\begin{aligned} a_{00}b_{00} &= c_{00} \\ a_{10}b_{00} + a_{00}b_{10} &= c_{10} \\ a_{01}b_{00} + a_{00}b_{01} &= c_{01} \\ a_{00}b_{11} + a_{01}b_{10} + a_{10}b_{01} + a_{11}b_{00} &= c_{11} \\ a_{20}b_{00} + a_{10}b_{10} + a_{00}b_{20} &= c_{20} \\ a_{02}b_{00} + a_{01}b_{01} + a_{00}b_{02} &= c_{02} \\ &\ldots \end{aligned} \tag{122}$$

The algebra is straightforward, as one notices that each c_{ij} is the sum of all configurations in (p, q, r, s) of the products $a_{pq}b_{rs}$, where $p + r = i$ and $q + s = j$.

If a deconvolution is to be carried out (i.e. to determine the original image after it was convolved with a filter function), the coefficients b_{ij} and c_{ij} are known. Indeed, the coefficients are respectively the 2D moments of the filter function and those of the degraded image. This is a straightforward generalization to two independent variables of similar results in one dimension. The unknown coefficients a_{ij} are determined by the linear system in Eq.(122). Their solution can be obtained rapidly with Gaussian elimination. Once the moments a_{ij} of the original image are determined, the image can be reconstructed by approximating the resulting hyperdistribution with an HR expansion, much in the same way it is done in one dimension. If, on the other hand, one is simply looking for the convolution inverse to f, that is to determine g such that $f * g = \delta(x, y)$, then $c_{00} = 1$ and $c_{nm} = 0$ for all $(n, m) \neq (0, 0)$; and the coefficients b_{ij} are once again easily determined.

3 Hermite-Rodriguez Wavelet Analysis

3.1 Introduction

Wavelets are an exciting new technique for solving difficult problems in mathematics, physics, and engineering. Applications are as diverse as seismic exploration, data compression of digitized signals and images, the detection of submarines, and improvements in CAT scans and other medical imaging technologies. Wavelets allow complex information such as speech, music, photographs, or video images, to be broken down into fundamental building blocks -the wavelets-, and subsequently reconstructed with high precision.

A recent article in Business Week magazine (Feb. 3 1992), entitled *'Wavelets are causing ripples everywhere'*, emphasizes:

> Catching the Wavelet– A dazzling theory of mathematics is leading to advances in: data compression (Wavelets identify key features of an image, allowing engineers to reconstruct a photo with only a tiny fraction of the information in the original), scientific calculation (Wavelets may help decipher phenomena such as turbulence– leading, for example, to improved aircraft designs), signal analysis (the new method makes it easier for military radar to spot hidden tanks), and medical imaging (scientists are using Wavelets to produce magnetic resonance pictures of the body faster and more accurately).

The decomposition of signals and images into a set of fundamental building blocks is not however a new concept. Fourier analysis is such a decomposition. Fourier methods decompose a signal into a series of sines and cosines. The discrete set of fundamental frequencies (the "building blocks") and their corresponding amplitude completely describe the initial signal, provided some care is taken in the discrete sampling procedure. Various properties (e.g. the convolution property) combine to make Fourier analysis a very attractive and useful mathematical tool.

Wavelets are however an entirely new set of building blocks that present a number of new features. One advantage of the new wavelet theory lies in its ability to zoom in on details –much like a camera with a zoom lens– by probing signals at different scales of resolution. Yet another advantage lies in the digital processing of signals and images: by processing the elementary building blocks, instead of the signal or the image, the processing part is simplified, and new insights are gained.

A novel parametric expansion, called *Hermite-Rodriguez* expansion, has been introduced as functional approximation to a new class of highly singular functions: hyperdistributions. The approximation is similar to the way a Gaussian of narrowing width is a Lighthill approximation[10] to the Dirac delta "function" $\delta(x)$; that is, the limit of the sequence of Gaussians of narrowing width is the Dirac delta function:

$$\lim_{\lambda \to 0} \frac{e^{-x^2/\lambda^2}}{\sqrt{\pi}\lambda} = \delta(x) \tag{123}$$

We will now prove that the underlying basis functions which generate Hermite-Rodriguez expansions are in fact wavelet functions[27].

The HR expansion of a function $f(x)$ is written as:

$$f(x) = \sum_{n=0}^{\infty} a_n \nabla^n \delta_\lambda(x) \tag{124}$$

The Rodriguez formula for Hermite polynomials[2] is then used to obtain:

$$f(x) = \sum_{n=0}^{\infty} a_n \, H_n(x/\lambda) \, \delta_\lambda(x) \tag{125}$$

with

$$a_n = \frac{1}{2^n n!} \int_{-\infty}^{+\infty} f(x)\, H_n(x/\lambda)\, dx \tag{126}$$

where $H_n(x)$ denotes the Hermite Polynomial in x of order n. We now redistribute 2^n and $n!$ in the following way:

$$f(x) = \sum_{n=0}^{\infty} a_n\, Hd_n^{\lambda}(x)\, \delta_{\lambda}(x) \quad \text{where} \quad a_n = \int_{-\infty}^{+\infty} Hd_n^{\lambda}(y) f(y) dy \tag{127}$$

where

$$Hd_n^{\lambda}(x) = \frac{H_n(x/\lambda)}{2^{n/2}\sqrt{n!}} \tag{128}$$

By using a well-known[3] bound on Hermite Polynomials, it can be shown that the Hd Polynomials are bounded in n:

$$\mid Hd_n^{\lambda}(x) \mid < k\, e^{x^2/2\lambda^2} \quad \text{where} \quad k \approx 1.086435 \tag{129}$$

Furthermore, the Hd Polynomials are orthogonal in the appropriate L^2 space, with

$$< Hd_n^{\lambda}, Hd_m^{\lambda} > = \int_{-\infty}^{+\infty} Hd_n^{\lambda}(x)\, Hd_m^{\lambda}(x)\, \frac{e^{-x^2/\lambda^2}}{\sqrt{\pi}\lambda}\, dx = \sqrt{\pi}\delta_{nm} \tag{130}$$

where δ_{nm} denotes Kronecker's delta; the generative recursion relation for Hd Polynomials becomes:

$$Hd_n^{\lambda}(x) = \sqrt{\frac{2}{n}}\frac{x}{\lambda}\, Hd_{n-1}^{\lambda}(x) - \sqrt{\frac{n-1}{n}}\, Hd_{n-2}^{\lambda}(x)$$

$$\text{with} \quad Hd_0^{\lambda}(x) = 1 \quad \text{and} \quad Hd_1^{\lambda}(x) = \frac{\sqrt{2}}{\lambda}x \tag{131}$$

The following "analyzing"[27] smooth point-functions are now introduced:

$$W_n^{\lambda}(x) = Hd_n^{\lambda}(x)\, \delta_{\lambda}(x) \tag{132}$$

These functions present the following properties:

1. $W_n^{\lambda}(x)$ is a function of class C^n.
2. $W_n^{\lambda}(x)$ has a rapid Gaussian-like decay at infinity. In other words, the support of the function is said to be *essentially* compact.

$W_n^{\lambda}(x)$ is called the Hermite-Rodriguez (HR) *wavelet*[27] of order n and weight λ. The Hermite-Rodriguez (HR) expansion of a signal $f(x)$ is given by

$$f(x) = \sum_{n=0}^{\infty} a_n W_n^{\lambda}(x) \quad \text{where} \quad a_n = \int_{-\infty}^{+\infty} Hd_n^{\lambda}(x) f(x)\, dx \tag{133}$$

A sample of the set of HR wavelets $\{W_n^\lambda(x)\}_n$ with $\lambda = 2$ are displayed in figure 5.

In a similar fashion, the HR expansion of an image $I(x,y)$ with weights λ and μ is given by

$$I(x,y) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} a_{nm} W_{nm}^{(\lambda,\mu)}(x,y) \quad \text{where} \quad a_{nm} = \iint_{-\infty}^{+\infty} Hd_n^\lambda(x) Hd_m^\mu(y) I(x,y)\, dxdy \tag{134}$$

with a similar convergence condition on weights (λ,μ), and with the 2D HR wavelet function defined as the tensor product of the two 1D independent HR wavelet functions:

$$W_{nm}^{(\lambda,\mu)}(x,y) = W_n^\lambda(x) W_m^\mu(y) \tag{135}$$

A sample of the positive part of the set $\{W_{nn}^{(\lambda,\lambda)}(x,y)\}_{n,n}$ for $\lambda = 2$ is displayed in figure 6, with grey scales replacing the ordinate values.

Hermite-Rodriguez wavelet expansions preserve the classical properties derivable for orthogonal polynomial expansions (as opposed to power series), but also add an important new feature: the weight parameter λ. The convergence condition derived previously from Christoffel-Darboux theory holds generally on a semi-infinite range of λ. This freedom can be utilized to optimize the rate of convergence of the HR wavelet expansion. HR wavelet expansions are also found to be *robust* with respect to numerical errors on the HR moment coefficients. In other words, if small errors are committed on the exact values of the moment coefficients (for example by quantizing the moments from double-precision reals to reals with one or two decimal places, or even to integers), the convergence properties of the resulting wavelet expansion do not suffer.

3.2 Hermite-Rodriguez Wavelet Expansion of a Gaussian

The Hermite-Rodriguez wavelet expansion of a Gaussian of unit width is given by:

$$e^{-x^2}/\sqrt{\pi} = \sum_{n=0}^{\infty} \frac{(1-\lambda^2)^n}{2^{3n/2}\sqrt{n!}} W_{2n}^\lambda(x) \tag{136}$$

A proof based on linear operator relations is as follows. It can be checked by differentiation in t and x that

$$\frac{e^{-x^2/4t}}{\sqrt{4\pi t}} = e^{t\nabla^2}\delta(x) \tag{137}$$

Letting

$$4t = \lambda^2 \tag{138}$$

We can write

$$\frac{e^{-x^2/\lambda^2}}{\sqrt{\pi}\lambda} = e^{\lambda^2\nabla^2/4}\delta(x) = \delta_\lambda(x) \tag{139}$$

Consider the identity

$$\begin{aligned} e^{\nabla^2/4}\delta(x) &= e^{\frac{1-\lambda^2}{4}\nabla^2} e^{\lambda^2\nabla^2/4}\delta(x) \\ &= e^{\frac{1-\lambda^2}{4}\nabla^2}(e^{\lambda^2\nabla^2/4}\delta(x)) \\ &= e^{\frac{1-\lambda^2}{4}\nabla^2}\delta_\lambda(x) \end{aligned} \tag{140}$$

Expanding the right-hand side,

$$e^{\nabla^2/4}\delta(x) = e^{-x^2}/\sqrt{\pi} = \sum_{n=0} \frac{(1-\lambda^2)^n}{2^{3n/2}\sqrt{n!}} W_{2n}^{\lambda}(x) \tag{141}$$

This method of proof provides a useful check on the numerical calculations of HR moments.

3.3 Hermite-Rodriguez Wavelet Expansions in One Dimension

In this section, the convergence properties of Hermite-Rodriguez wavelet expansions are investigated. To this end, we introduce a set of 80 artificial "landscapes", the first 16 of which are displayed in figure 7. These landscapes are generated by summing 5 Gaussians whose widths and offsets are generated randomly in the range [0.1,1] for the widths, and [-3,3] for the offsets. The landscapes are plotted roughly within the support of three standard deviations, ie. in the range [-6,6]. In other words, if L_n denotes the n^{th} landscape, then:

$$L_n(x) = \sum_{i=1}^{5} \delta_{\lambda_i}(x - x_i) \tag{142}$$

where $\delta_\lambda(x)$ denotes the Gaussian of width λ, and (λ_i, x_i) are computer-generated pseudo-random variables with respective ranges [0.1,1] and [-3,3]. All landscapes are then rescaled to unit height. Each landscape is sampled by 501 datapoints (501 numbers in the range [-6,6]).

If all 5 widths or all 5 offsets are close together in their respective values, the associated landscape is labeled "realistic" (ie. the landscape resembles a realistic mountain or mountain range), while if on the contrary, the 5 widths and offsets are spread out in range, the generated landscape is labeled "futuristic".

Although Hermite polynomials of high order are known to diverge polynomially (x^n, "whipping-tail") as their order n increases, the Hermite-Rodriguez wavelets of high order are well-behaved, and decreasing in modulus. More specifically, a classical bound of Hermite Polynomials[3] yields:

$$| W_n^{\lambda}(x) | < k \frac{e^{-x^2/2\lambda^2}}{\sqrt{\pi}\lambda} \quad \text{where} \quad k \approx 1.086436 \tag{143}$$

Figure 8 displays the HR wavelets of orders 1000 and 10000.

In Figure 9, a shifted Gaussian (centered at $x = 2$), of width $\lambda = 1/2$, renormalized to unit height, denoted by $L_0(x)$, is approximated with a Hermite-Rodriguez wavelet expansion $A_\lambda(x)$ of weight $\lambda = 1$, with N^{th} partial sum:

$$A_\lambda^N(x) = \sum_{i=0}^{N} a_i W_i^\lambda(x) \quad \text{where} \quad a_i = \int_{-6}^{+6} L_0(x) Hd_i^\lambda(x)\, dx \tag{144}$$

where

$$W_i^\lambda(x) = Hd_i^\lambda(x)\delta_\lambda(x) \quad \text{and} \quad Hd_i^\lambda(x) = \frac{H_i(x/\lambda)}{2^{n/2}\sqrt{n!}} \tag{145}$$

and $H_i(x)$ denotes the Hermite polynomial of order i. The expansion is seen to converge. The 40^{th} partial sum already closely approximates the original Gaussian.

In figure 10, a futuristic-looking landscape (landscape #68 and denoted by $L_{68}(x)$) is approximated with a Hermite-Rodriguez wavelet expansion of width $\lambda = 1$. The 300^{th} partial sum closely approximates the original landscape. A high number of partial sums is required for the capture of high frequency contents.

In figures 11 and 12, a less futuristic-looking (but still futuristic) landscape (landscape #71 and denoted by $L_{71}(x)$) is approximated with a Hermite-Rodriguez wavelet expansion for two different values of the weight λ: 1.0 and 1.15 respectively. The 140^{th} partial sum for $\lambda = 1.15$ quickly approximates the original landscape, while the wavelet expansion for $\lambda = 1.0$ has a much slower convergence rate. This free parameter, can thus be utilized to improve rate of convergence. Figures 13 through 17 emphasize the critical role of the weight parameter λ in improving the convergence of Hermite-Rodriguez wavelet expansions. Figure 18 displays the robustness of HR expansions, when HR moments are rounded off. The rate of convergence of the quantized wavelet expansion is only very slightly modified.

It can be observed that landscapes with an increasing number of higher frequencies require an increasing number of partial sums for accurate HR wavelet expansion representations. This was expected for HR wavelet expansions, which are of global nature (moment expansions vs. Taylor expansions), and capture global behavior very fast, while they are slower in capturing sharply localized features.

3.4 Hermite-Rodriguez Wavelet Expansions in Two Dimensions

In a similar fashion, the HR wavelet expansion of an image $I(x,y)$ with weights λ and μ is given by

$$I(x,y) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} a_{nm} W_{nm}^{(\lambda,\mu)}(x,y) \quad \text{where} \quad a_{nm} = \iint_{-\infty}^{+\infty} Hd_n^{\lambda}(x) Hd_m^{\mu}(y) I(x,y) dx dy \tag{146}$$

with the 2D HR Wavelet function defined as the tensor product of the two 1D independent HR Wavelet functions:

$$W_{nm}^{(\lambda,\mu)}(x,y) = W_n^{\lambda}(x) W_m^{\mu}(y) \tag{147}$$

A sample of the positive part of the set $\{W_{nn}^{(\lambda,\lambda)}(x,y)\}_{n,n}$ for $\lambda = \mu = 2$ is displayed in figure 6, with grey scales replacing the ordinate values.

A set of 80 artificial "nebulae", the first 16 of which are displayed in figure 19. These nebulae are generated by summing 5 2D Gaussians whose widths and offsets are generated randomly in the range [0.2,1]×[0.2,1] for the widths, and [-2,2]×[-2,2] for the offsets. The nebulae are plotted within the range [-6,6]×[-6,6]. In other words, if N_n denotes the n^{th} nebula, we have:

$$N_n(x) = \sum_{i=1}^{5} \delta_{\lambda_i,\mu_i}(x - x_i, y - y_i) \tag{148}$$

where $\delta_{\lambda_i,\mu_i}(x,y)$ denotes the 2D Gaussian of width λ_i and μ_i, which is the tensor product of two 1D Gaussians:

$$\delta_{\lambda_i,\mu_i}(x,y) = \frac{e^{-\frac{x^2}{\lambda_i^2}}}{\sqrt{\pi}\lambda_i} \frac{e^{-\frac{y^2}{\mu_i^2}}}{\sqrt{\pi}\mu_i} \tag{149}$$

and (λ_i, μ_i) and (x_i, y_i) are computer-generated pseudo-random variables with respective ranges [0.2,1] and [-2,2]. All nebulae are then rescaled to unit height. Each nebula is sampled by 64×64 datapoints.

4 Green's Function for the Antidiffusion Equation

In this section, we repeat in more detail two proofs given in sections 1 and 2. Namely that Fourier transforms do not provide us with a Green's function for the antidiffusion equation (i.e. a convolution inverse of a gaussian), but that, in contrast, there is a simple hyperdistribution representation for that same Green's function. We then attempt to reconstruct images corrupted with gaussian blur by letting the blurred images evolve freely as initial conditions of the antidiffusion equation.

4.1 Antidiffusion and Fourier Transforms

Consider the Gaussian filter $G(x,t)$:

$$G(x,t) = \frac{e^{-\frac{x^2}{4t}}}{\sqrt{4\pi t}} \tag{150}$$

We suppress the spatial variable x if confusion does not arise, and denote $G(x,t)$ by $G(t)$. We have shown in section 1.2 that an attempt to evaluate an explicit representation of the convolution inverse of $G(t)$ from Fourier transforms yields:

$$F(a,x,2t) = \frac{1}{2\pi}\int_{-a}^{+a} e^{-ikx+k^2t}\,dk \tag{151}$$

$$= \frac{e^{\frac{x^2}{4t}}}{4i\sqrt{\pi t}}[\mathrm{erf}(i\sqrt{t}\,a + \frac{x}{2\sqrt{t}}) + \mathrm{erf}(i\sqrt{t}\,a - \frac{x}{2\sqrt{t}})] \tag{152}$$

where $\mathrm{erf}(x)$ denotes the error function[3]. Proceeding formally, $Inv[G(t)] = \lim_{a\to\infty} F(a,x,2t)$, which was shown to diverge.

$G(t)$ is a smooth point function for nonegative values of t. By contrast, $Inv[G(t)]$ is not even a distribution, since it does not exhibit weak convergence[4]-[11]. Specifically, its integral over the Gaussian test function $e^{-\frac{x^2}{4t}}$ diverges. Proceeding with the use of Fubini's theorem[25] on the interchange of the limit and integral sign,

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{16t}} \lim_{a\to\infty} F(a,x,2t)\,dx = \lim_{a\to\infty}\int_{-a}^{+a} e^{k^2t}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{4t}-ikx}\,dx\,dk = \lim_{a\to\infty} 4a\sqrt{\pi t} \to +\infty \tag{153}$$

In other words, the convolution inverse of a Gaussian filter as given by Fourier transform methods is not a smooth point-function.

4.2 Antidiffusion and Hyperdistributions

In the previous paragraph, it was proven that $Inv[G(t)]$, as obtained by Fourier transform methods, is ill-defined. By contrast, we can give an explicit representation of $Inv[G(t)]$ using hyperdistributions. The Green's function for the diffusion equation[21][22] can be written[23] as

$$G(x,t) = e^{t\nabla^2}\delta(x) \tag{154}$$

where δ denotes the Dirac delta function, and which, for $t \geq 0$, converges to the well known Gaussian given above in Eq.(150). The function G, also called the *initial condition Green's function*[24], has the properties

$$\frac{\partial}{\partial t}G(x,t) = \nabla^2 G(x,t) \quad , \quad G(x,0) = \delta(x) \tag{155}$$

We note that the two infinite series of distributions:

$$e^{\pm t\nabla^2}\delta(x) = \sum_{n=0}^{\infty} \frac{1}{n!}(\pm\, t\nabla^2)^n \delta(x) \tag{156}$$

are hyperdistributions. Furthermore, the expressions are inverse to each other under convolution, as can be checked by the Bochner algebra for hyperdistributions. The convolution inverse $Inv[G(t)]$ of the Gaussian is now readily obtained using Eq.(154). Thus

$$\delta(x) = (e^{t\nabla^2}\delta(x)) * Inv[G(t)] = e^{t\nabla^2}(\delta * Inv[G(t)])(x) = e^{t\nabla^2} Inv[G(t)] \tag{157}$$

from which we conclude that

$$Inv[G(t)] = e^{-t\nabla^2}\delta(x) \tag{158}$$

We note the relation of the convolution inverse $Inv[G(t)]$ to the time reflected solution of the diffusion equation:

$$Inv[G(t)] = G(-t) = G^{\mathcal{T}}(t) \tag{159}$$

where $\mathcal{T}$ denotes "time reflected". The quantity $G^{\mathcal{T}}(t)$ satisfies the equation:

$$\frac{\partial}{\partial t} G^{\mathcal{T}}(t) = -\nabla^2 G^{\mathcal{T}}(t) \quad , \quad G^{\mathcal{T}}(0) = \delta(x) \tag{160}$$

and therefore $Inv[G(t)]$ is the fundamental solution of the antidiffusion equation (time-reversed diffusion). The sequence $\{F(a, x, 2t)\}_a$ plotted in figure 4 can be shown to define the same Hyperdistribution given by Eqs.(156) and (158). The sequence in Eq.(156) is analogous to the sequence that defines $\delta(x)$ in terms of narrowing Gaussians. Both G and $Inv[G]$ have analytic power series in wave-number space with Eq.(156) as Fourier images. This is not accidental; in fact, hyperdistributions have the algebraic properties of formal power series, since the Fourier transform of hyperdistributions is a power series in the wave-number k.

4.3 Antidiffusing Images Corrupted with Gaussian Blur

We have shown in the previous section that the Bochner algebra for deconvolving hyperdistributions leads to a theoretically possible reconstruction of waveforms blurred by a Gaussian filter by antidiffusing the blurred waveform for an appropriate duration of time. In this section, we will go through the proof in more detail.

Suppose an image is blurred by convolution with a two-dimensional Gaussian filter of width λ. The convolution inverse to the Gaussian filter is obtained with the help

of the Bochner algebra for hyperdistributions. The 2D Hermite-Rodriguez moments of a normalized Gaussian filter $\delta_\lambda(x,y)$ of width λ are given exactly by

$$a_{nm} = \iint_{-\infty}^{+\infty} \frac{e^{-\frac{x^2+y^2}{\lambda^2}}}{\sqrt{\pi}\lambda} x^n y^m \, dx\, dy = \begin{cases} \frac{n!m!}{(\frac{n}{2})!(\frac{m}{2})!} \frac{\lambda^{2(n+m)}}{2^{3(n+m)/2}} & \text{if } n \text{ and } m \text{ are even.} \\ 0 & \text{otherwise.} \end{cases} \tag{161}$$

Thus, the hyperdistribution expansion of the Gaussian filter:

$$\begin{aligned} \delta_\lambda(x,y) &= \sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty} \frac{(-1)^n}{n!}\frac{(-1)^m}{m!} a_{nm} \nabla_x^n \delta(x) \nabla_y^m \delta(y) \\ &= \sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty} (\frac{\lambda^4}{8})^{n+m} \frac{1}{n!m!} \nabla_x^{2n} \nabla_y^{2m} \delta(x)\delta(y) \end{aligned} \tag{162}$$

which can be separated in the form

$$\delta_\lambda(x,y) = \left\{ \sum_{n=0}^{+\infty} (\frac{\lambda^4}{8})^n \frac{1}{n!} \nabla_x^{2n} \delta(x) \right\} \left\{ \sum_{m=0}^{+\infty} (\frac{\lambda^4}{8})^m \frac{1}{m!} \nabla_y^{2m} \delta(y) \right\} \tag{163}$$

The above double sum can be written in simple operator form as

$$\delta_\lambda(x,y) = e^{\frac{\lambda^4}{8}\nabla_x^2}\delta(x) e^{\frac{\lambda^4}{8}\nabla_y^2}\delta(y) = e^{\frac{\lambda^4}{8}\nabla^2}\delta(x)\delta(y) \tag{164}$$

The convolution inverse $Inv[\delta_\lambda]$ of δ_λ is now easily determined through the 2D Bochner Algebra. More specifically, it represents a special simplifying case where the 2D algebra reduces to the product of two 1D Bochner algebras. That is because the above hyperdistribution of two independent variables breaks down into the product of two hyperdistributions of a single variable (respectively x and y), as can be seen in Eq.(163). This is not surprising; a two-dimensional Gaussian is separable in its two independent variables. Thus, for all such "radial" filters, the Bochner algebra reduces to the product of two 1D algebras:

$$c_{nm} = \left\{ \sum_{p=0}^{n} a^1_{n-p} b^1_p \right\} \left\{ \sum_{q=0}^{m} a^2_{n-p} b^2_p \right\} \tag{165}$$

where the superscripts ^{1} and ^{2} refer to the respective independent variables x and y.

The convolution inverse to δ_λ is then easily determined (see appendix A) to be

$$Inv[\delta_\lambda](x,y) = \sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty} (-\frac{\lambda^4}{8})^{n+m} \frac{1}{n!m!} \nabla_x^{2n} \nabla_y^{2m} \delta(x,y) \tag{166}$$

which can be written again as

$$Inv[\delta_\lambda](x,y) = e^{-\frac{\lambda^4}{8}\nabla^2}\delta(x,y) \tag{167}$$

Thus, filtering an image $I(x,y)$ with $\delta_\lambda(x,y)$ yields the blurred image $J(x,y)$ where

$$J(x,y) = \delta_\lambda(x,y) * I(x,y) = e^{\frac{\lambda^4}{8}\nabla^2}\delta(x,y) * I(x,y) = e^{\frac{\lambda^4}{8}\nabla^2} I(x,y) \tag{168}$$

And finally, reconstructing the image $I(x,y)$:

$$I(x,y) = J(x,y) * Inv[\delta_\lambda](x,y) = e^{-\frac{\lambda^4}{8}\nabla^2}\delta(x,y) * J(x,y) = e^{-\frac{\lambda^4}{8}\nabla^2} J(x,y) \tag{169}$$

The use of linear operators allows simple representations for all the quantities involved.

Eq.(169) can then be further simplified with the help of Taylor expansions in the following way. Write down $\lambda^4/8$ as a finite sum of equally small quantities. Then $\lambda^4/8 = \sum_{k=0}^{n} \Delta_k$ and

$$I(x,y) = e^{-\frac{\lambda^4}{8}\nabla^2} J(x,y) = e^{-\sum_{k=0}^{n}\Delta_k\nabla^2} J(x,y) = \prod^{n} \left\{ e^{-\Delta_k\nabla^2} J(x,y) \right\} \tag{170}$$

And since Δ_k is small, $e^{-\Delta_k\nabla^2}$ can be approximated to first order by its Taylor expansion:

$$e^{-\Delta_k\nabla^2} = 1 - \Delta_k\nabla^2 + \mathcal{O}(\Delta_k^2) \tag{171}$$

Eq.(170) can thus be approximated to first order by:

$$I(x,y) = \prod^{n} \left\{ J(x,y) - \Delta\nabla^2 J(x,y) \right\} \tag{172}$$

where the subscript k in Δ has been dropped. It can now be shown that the original image can be recovered by running a *time-reversed diffusion* with the degraded image J as initial condition. The normalized time-reversed diffusion equation (or antidiffusion) can be written as

$$\frac{\partial}{\partial t} J + \nabla^2 J = 0 \tag{173}$$

and can be forward-discretized in time as

$$\frac{J_{t+\Delta t} - J_t}{\Delta} = -\nabla^2 J_t \tag{174}$$

or equivalently as:

$$J_{t+\Delta t} = J_t - \Delta\nabla^2 J_t \tag{175}$$

Equations (172) and (175) are one and the same, and therefore the original image can theoretically be reconstructed by applying a time-reversed diffusion on the degraded image with a very small time-step (so that Eq.(171) holds to first order). Higher-order corrections to our approximation can be implemented by expanding the exponential in Eq.(170) to higher order.

We now discretize the 2+1 antidiffusion/diffusion equation with diffusivity $\pm\nu$:

$$\frac{\partial u}{\partial t} \pm \nu \nabla^2 u = 0 \tag{176}$$

in the following way. The infinite plane is sampled at locations on a square grid:

$$u \longrightarrow u_{ij}^k \tag{177}$$

where the subscripts i and j represent the spatial location of a mesh point, and the superscript k represents the time index. The continuous operators ∇^2 and $\partial/\partial t$ are replaced by the discrete approximations:

$$\frac{\partial u_{ij}^k}{\partial t} \longrightarrow \frac{u_{ij}^{k+1} - u_{ij}^k}{\delta_t} \tag{178}$$

where δ_t represents the discretized time-step, (this is a forward-in-time discretization scheme, with an error of the order of δ_t^2), and

$$\nabla^2 u_{ij}^k \longrightarrow \frac{\alpha \sum u^k|_\diamond + \beta \sum u^k|_\square + \gamma u_{ij}^k}{(\alpha+\beta)\delta_x^2} \tag{179}$$

where δ_x represents the discretized space-gap, and

$$\gamma = -4(\alpha+\beta) \tag{180}$$

$\sum u^k|_\diamond$ represents the sum of all immediate neighbors of u_{ij}^k forming a diamond pattern centered at u_{ij}^k, and $\sum u^k|_\square$ represents the sum of all immediate neighbors of u_{ij}^k forming a square pattern centered at u_{ij}^k (this is a central-in-space discretization scheme, with an error of the order of δ_x^4). The configuration of neighbors forming a diamond pattern is called the *diamond stencil*, while the configuration of neighbors forming a square pattern is called the *square stencil*. This central-in-space discretization of the continuous operator ∇^2 is carried out in appendix B. How is the ratio α/β to be chosen[28]? In other words, what is the weighting of the diagonal and the square neighbors in the discretization of the operator ∇^2? It is found that the best scheme involves weighing diamond and square stencils based on their mesh-distance from the central grid-point (square neighbors are by a factor $\sqrt{2}$ further apart from the central grid-point than diamond neighbors):

$$\alpha/\beta = \sqrt{2} \tag{181}$$

The numerical values chosen in the discretization are the following. δ_t must be small for the approximation in Eq(171) to hold to second order, and δ_x must be small for the approximation in Eq.(179) to hold to fourth order. More specifically, we have chosen $\nu = 1\,L^2/T$, $\delta_x = 12/65\,L$, $\delta_t = 1.7 \times 10^{-6}\,T$, where L is a unit of length,

and T is a unit of time. The values of t specified in the figures have been rescaled to higher values for ease of comparison. In order to relate these diffusion durations to the duration of time T responsible for the diffusion $e^{T\nabla^2}I(x,y)$ of the image $I(x,y)$, divide t by the factor 0.0085.

In figures 20 through 23, images of the letters "T" and "E" are diffused, and a reconstruction is attempted by antidiffusing the blurred images. If the diffusion duration is too long, the attempted reconstruction fails.

In figure 24, an initial pattern in the shape of a spiral is diffused, and a reconstruction is attempted for increasing amounts of diffusion duration. Once again, it is observed that if the diffusion duration is too long, the reconstruction fails.

In figure 25, the effect of noise on reconstruction by antidiffusion is analyzed. An intial pattern in the shape of a letter "T" is diffused for a certain amount of time. A 2D field of pseudo-white noise is generated with a random-number-generator. That field is subsequently added onto the diffused pattern at a proportion given by the Signal-to-Noise Ratio (SNR). The resulting pattern is subsequently antidiffused for the appropriate amount of time, and the resulting reconstructions are displayed.

We conclude by noticing that the reconstruction by antidiffusion of blurred images is an unstable algorithm, with the noise-content of the image (or numerical round-off errors) as the source of the instability. Each mode is exponentially amplified in magnitude with antidiffusion. The Signal-to-Noise Ratio thus plays a critical perturbative role. The reconstruction process breaks down if the width of the Gaussian filter is too wide (i.e. if the diffusion and thus antidiffusion durations are too long: figures 21, 23, and 24), or if the Signal-to-Noise Ratio is too low[5] (figure 25).

In the next section, we will approximate hyperdistribution sums with Hermite-Rodriguez Wavelets, and once again attempt to reconstruct signals and images corrupted with gaussian blur, in the difficult case of low Signal-to-Noise Ratios. We will then generalize my results to include the case of non-gaussian blur.

5 Deblurring with Hermite-Rodriguez Wavelets

5.1 Introduction

In the previous section, we have concluded that deblurring by discretization of the antidiffusion equation is a chaotic process, since numerical errors increase exponentially with time. In other words, antidiffusion is extremely sensitive to noise, and an

[5] round-off error in the computation is equivalent to noise added onto the blurred image. The computation is implemented with double-precision fortran real variables, corresponding to a phenomenal 10^{15} SNR, as long as the blurred image is free of any other kind of noise

extremely high Signal-to-Noise Ratio (SNR) is necessary for reconstructing a blurred image via antidiffusion[14]. In this section, we modify the antidiffusion algorithm by expanding the blurred waveform in the form of a Hermite-Rodriguez wavelet expansion. The antidiffusion is then carried out on the building blocks of the expansion –the HR wavelets–, as opposed to the waveform as a whole. Hermite-Rodriguez wavelets are known exactly, and can thus be more successfully antidiffused when the blurred waveform is corrupted by noise. The noise enters effectively only on the Hermite-Rodriguez moments, which are constant and remain unchanged by antidiffusion. This deblurring procedure is found to be more robust with respect to noise. Artificial 1D signals ("landscapes") and artificial 2D astronomical images ("nebulae"), which have been blurred by a diffusion process and subsequently corrupted with white noise, are successfully reconstructed.

5.2 Hermite-Rodriguez Wavelets and Antidiffusion

The fundamental solution of the diffusion equation[21][22] at time t can be written[23] as

$$G(x,t) = e^{t\nabla^2}\delta(x) \tag{182}$$

where δ denotes the Dirac delta function, and which, for $t \geq 0$, converges to the Gaussian form:

$$G(x,t) = \frac{e^{-x^2/4t}}{2\sqrt{\pi t}} \tag{183}$$

The function G, also called the *initial condition Green's function*[24], has the properties

$$\frac{\partial}{\partial t}G(x,t) = \nabla^2 G(x,t) \quad , \quad G(x,0) = \delta(x) \tag{184}$$

It has been proven that the two hyperdistributions defined by

$$e^{\pm t\nabla^2}\delta(x) = \sum_{n=0}^{\infty}\frac{1}{n!}(\pm\, t\nabla^2)^n\delta(x) \tag{185}$$

are inverse to each other under convolution. That is, the convolution $e^{-t\nabla^2}\delta(x) * e^{+t\nabla^2}\delta(x)$ yields the Dirac delta function $\delta(x)$. Consequently, the fundamental solution of the antidiffusion equation

$$\frac{\partial}{\partial t}G(x,t) + \nabla^2 G(x,t) = 0 \tag{186}$$

can formally be written as $e^{-t\nabla^2}\delta(x)$.

The set of HR Wavelets $\{W_n^\lambda(x)\}_\lambda$ can be shown to be a globally invariant subspace (in λ) of the operators $e^{\pm t\nabla^2}$ (see appendix C). In other words, it can be proven that,

diffusing the HR wavelet of order n and weight λ_1 for a duration of time t yields a rescaled HR wavelet of same order but *higher* weight λ_2:

$$e^{+t\nabla^2} W_n^{\lambda_1}(x) = \left(\frac{\lambda_1}{\lambda_2}\right)^n W_n^{\lambda_2}(x) \tag{187}$$

where $\lambda_2 = \sqrt{\lambda_1^2 + 4t}$. Consequently, the above statement can be time-reversed, to show that the antidiffusion of the HR wavelet of order n and weight λ_2 for a duration of time t yields a rescaled HR wavelet of same order but *lesser* weight λ_1:

$$e^{-t\nabla^2} W_n^{\lambda_2}(x) = \left(\frac{\lambda_2}{\lambda_1}\right)^n W_n^{\lambda_1}(x) \tag{188}$$

This property is utilized in the following section.

5.3 Wavelet Reconstruction of Diffused Signals and Images (Gaussian Blur)

A one-dimensional signal $S(x)$ is diffused for a duration of time t. The resulting blurred[6] signal is denoted as $\tilde{S}(x)$. The goal is to reconstruct $S(x)$ by antidiffusing $\tilde{S}(x)$ for the same duration of time t. To this end, we expand $\tilde{S}(x)$ as the N^{th} partial sum of an HR expansion of weight λ_2:

$$\tilde{S}(x) \approx \sum_{n=0}^{N} \tilde{a}_n W_n^{\lambda_2}(x) \tag{189}$$

where $\{\tilde{a}_n\}_n$ are the Hermite moments of the blurred signal $\tilde{S}(x)$, weighted by λ_2. The partial sum above can be formally antidiffused by antidiffusing the constitutive blocks of the expansion, which are the HR wavelets: by formally antidiffusing for a duration of time t, we obtain:

$$e^{-t\nabla^2}\tilde{S}(x) \approx e^{-t\nabla^2}\{\sum_{n=0}^{N} \tilde{a}_n W_n^{\lambda_2}(x)\} = \sum_{n=0}^{N} \tilde{a}_n\, e^{-t\nabla^2}\{W_n^{\lambda_2}(x)\} = \sum_{n=0}^{N} \tilde{a}_n \left(\frac{\lambda_2}{\lambda_1}\right)^n W_n^{\lambda_1}(x) \tag{190}$$

with $\lambda_1 = \sqrt{\lambda_2^2 - 4t}$. The antidiffused signal can therefore be put in a form very similar to an HR expansion, this time of weight λ_1. However, the antidiffused signal in Eq.(190) is not an HR expansion, because of the extra factor $\left(\frac{\lambda_2}{\lambda_1}\right)^n$ multiplying the Hermite moments $\tilde{a}_n$. Consequently, a condition of the same type as in Eq.(114) does not guarantee the convergence of the antidiffused series. This matter will be treated in the next paragraph. The case of two-dimensional images is similar, since two-dimensional HR wavelets are the tensor product of two one-dimensional HR wavelets.

[6] in this section, we restrict ourselves to pure gaussian blur

As a result, the antidiffused HR expansion of an image can be put in a form similar to a 2D HR expansion, with the extra factor $\left(\frac{\lambda_2}{\lambda_1}\right)^{n+m}$ multiplying the Hermite moment of order $n \times m$.

Blurred signals and images (pure gaussian blur) have been *formally* reconstructed by antidiffusing their HR expansions. Since $e^{\pm t\nabla^2}$ is a linear operator, it is known[29] that[7] if both series

$$\tilde{S}(x) = \sum_{n=0}^{N} \tilde{a}_n W_n^{\lambda_2}(x) \tag{191}$$

$$R(x) = \sum_{n=0}^{N} \tilde{a}_n \left(\frac{\lambda_2}{\lambda_1}\right)^n W_n^{\lambda_1}(x) \tag{192}$$

are convergent, then the reconstructed signal is given by $R(x) = e^{-t\nabla^2}\tilde{S}(x)$, with $R(x)$ given by Eq.(192). However, there is no guarantee that the series $R(x)$ does converge.

More specifically, if the moments $\tilde{a}_n$ do not decrease to zero as fast as $\left(\frac{\lambda_2}{\lambda_1}\right)^n$ diverges as $n \rightarrow \infty$, then the series $R(x)$ diverges. This is in fact the typical case when the diffused waveforms are corrupted by instrumental, numerical, or computer round-off noise. To see how this happens, we decompose the Hermite moments $\tilde{a}_n$ for a diffused and noisy signal as the sum of the analytically *exact* (but unknown) moments $\tilde{b}_n$, and the residual error $\tilde{\epsilon}_n$:

$$\tilde{a}_n = \tilde{b}_n + \tilde{\epsilon}_n \tag{193}$$

The antidiffusion of the HR expansion with the analytically exact moments $\tilde{b}_n$ yields precisely the HR expansion which converges to the original signal $S(x)$:

$$S(x) = e^{-t\nabla^2}\left\{\sum_{n=0}^{\infty} \tilde{b}_n W_n^{\lambda_2}(x)\right\} = \sum_{n=0}^{\infty} \tilde{b}_n \left(\frac{\lambda_2}{\lambda_1}\right)^n W_n^{\lambda_1}(x) = \sum_{n=0}^{\infty} a_n W_n^{\lambda_1}(x) \tag{194}$$

On the other hand, the antidiffusion of the HR expansion with the residual error coefficients $\tilde{\epsilon}_n$ diverges:

$$\lim_{N\rightarrow\infty} \sum_{n=0}^{N} \tilde{\epsilon}_n \left(\frac{\lambda_2}{\lambda_1}\right)^n W_n^{\lambda_1}(x) = \infty \tag{195}$$

However, since usually $\tilde{\epsilon}_n \ll \tilde{b}_n$, the series in Eq.(195) will require a minimum amount of partial sums before the residual error can grow in modulus and overtake the exact solution as given by Eq.(194). The series in Eq.(195) is called partially convergent, or

[7] in one dimension, without loss of generality

asymptotic. Consequently, if the N^{th} partial sum $\sum_{n=0}^{N} a_n W_n^{\lambda_1}(x)$ is a good approximation of the original signal $S(x)$, and

$$\sum_{n=0}^{N} \tilde{\epsilon}_n \left(\frac{\lambda_2}{\lambda_1}\right)^n W_n^{\lambda_1}(x) \ll \sum_{n=0}^{N} a_n W_n^{\lambda_1}(x) \tag{196}$$

which usually implies that

$$\tilde{\epsilon}_N \left(\frac{\lambda_2}{\lambda_1}\right)^N \ll a_N \tag{197}$$

then the reconstruction

$$R(x) = \sum_{n=0}^{N} \tilde{a}_n \left(\frac{\lambda_2}{\lambda_1}\right)^n W_n^{\lambda_1}(x) \tag{198}$$

is successful, and $R(x) \approx S(x)$. If, on the other hand, the antidiffused residual error series grows too fast (e.g. if $\tilde{b}_n$ and $\tilde{\epsilon}_n$ are of the same order, which is the case if the Signal-to-Noise Ratio is too low), and higher-order partial sums are required for good approximations of the signals $\tilde{S}(x)$ and $S(x)$ (e.g. if the original signal comprises a large number of high-frequency components), then the reconstruction will fail. In other words, for lower order partial sums, antidiffusion can be implemented for longer durations of time, and consequently reconstruct severly blurred signals. If, on the other hand, good approximations of the signals $\tilde{S}(x)$ and $S(x)$ require partial sums of higher order, the amount of antidiffusion which can be performed on the blurred signal before instrumental and numerical errors overcome the calculation is limited. This limitation, as well as *successful* deconvolutions, are displayed in figure 26 for one-dimensional signals, and in figure 27 for two-dimensional images.

We now investigate whether it is possible to determine a priori the threshold in the order N for which the reconstruction by antidiffusion will fail. To this effect, We can utilize one of two classical constraints employed in signal and image reconstruction:

- **positivity**: the intensity of each pixel of the reconstructed waveform must be positive, since the original waveform is everywhere positive.
- **total intensity conservation**: the total intensity of the waveform is a conserved quantity under any diffusion/antidiffusion process (see appendix E). Consequently, it is required of the reconstructed waveform to conserve the total intensity of the blurred waveform.

At a certain threshold in the order N of the reconstruction, one of the two criteria above is destined to fail. At that point, the reconstruction will fail as well. That is because the diverging antidiffused residual error series in Eq.(195) overtakes in magnitude the value of the exact antidiffused series in Eq.(194). The upper limit in N for which the two criteria above are still verified will yield the order of the best possible reconstruction by antidiffusion of the HR expansion of the blurred waveform.

5.4 Application to Simulated Landscapes and Astronomical Images Corrupted with Gaussian Blur

In figures 26 and 27, a one-dimensional signal (landscape 11) and a two-dimensional image (nebula 62) are diffused for a certain amount of time. The diffused waveforms are subsequently corrupted with white noise obtained by a pseudo-random number generating algorithm. The one-dimensional signal is corrupted with 3% additive noise (i.e., the noise is added uniformly onto the diffused signal, at a proportion of 3% of the maximum intensity of the signal). The two-dimensional image is corrupted with 30% multiplicative noise (i.e. the noise is added uniformly onto the diffused image, pixel by pixel, at a varying proportion of 30% of the intensity of each pixel). Both the diffused and noisy signal and image are then approximated by a 1D and 2D HR expansion respectively. The HR expansions are subsequently antidiffused by antidiffusing the underlying HR wavelets. The antidiffusion of the 1D HR wavelet of order n and weight λ for an amount of time t is known exactly: it yields the HR wavelet of same order, of weight $\sqrt{\lambda^2 - 4t}$, rescaled by the factor $\{\lambda/(\sqrt{\lambda^2 - 4t})\}^n$. The antidiffusion is carried out on the blurred signal (for the same amount of time as the earlier diffusion duration) on two HR partial sums of different order, one of lower order, the other of higher order. Both reconstructions are displayed. The antidiffusion of the lower-order partial sum is successful, while the antidiffusion of the higher-order partial sum displays the diverging effects of noise on the Hermite moments. The antidiffusion is also carried out on the blurred image for two different HR partial sums. The same remarks apply. The tests of positivity and conservation of total intensity are carried out for the signal in figure 26 and the image in figure 27. The results are displayed in figure 28, together with a least-squares goodness-of-fit evaluation of the fidelity of the reconstructed waveform with the original waveform, plotted against the order N of the partial sum. The best reconstruction is located at the minimum of the least-squares curve, and this does correspond to the upper limit in the order N for which the reconstructed waveform is everywhere positive, and for which the total intensity remains equal to the total intensity of the blurred waveform.

It is noteworthy to point out that the noisy signals and images were *not* pre-processed for noise reduction before antidiffusion.

In figure 29, the blurring of the image in figure 27 is repeated, and a reconstruction of the original image by a direct discretization of the antidiffusion equation is attempted. In the case when the blurring is clean –that is, when the only source of noise on the diffused image is numerical round-off noise[8]–, the reconstruction is successful. In the case when the diffused image is corrupted by noise, the reconstruction is highly unsuccessful.

[8]the discretization is performed with double-precision real variables, bringing the Signal-to-Noise Ratio in this case to about 10^{15}.

The advantages of wavelet-based antidiffusion as an algorithm for deblurring signals and images are:

- The antidiffusion algorithm does not require any lengthy computation, since antidiffused wavelets are known exactly. This is in sharp contrast with the antidiffusion algorithm (directly involving a discretization of the antidiffusion equation, which has a very high computational load), or even Fourier transform methods (which require lengthy integral calculations).
- The algorithm is noise-robust, since noise enters only in the computation of the Hermite moments of the blurred waveform. In other words, noise is involved in an integration (the moment computation), which is a "smoothing" operation, as opposed to a differentiation process (e.g. direct discretization of the antidiffusion equation), which is a "roughening" operation. Nevertheless, noise or numerical errors on the Hermite moments eventually impair the reconstruction of the blurred waveform in the case where higher-order partial sums are necessary for accurate HR-expansion approximations of the blurred and original waveforms, as well as in the case of large diffusion durations.

5.5 * Wavelet Reconstruction of Signals and Images corrupted with non-Gaussian Blur

A one-dimensional signal $S(x)$ is blurred by convolution with a filter $F(x)$ that is not a gaussian (in other words, the blurring process is not equivalent to a diffusion in time, but rather the result of a more complicated process). The resulting blurred signal is denoted as $I(x)$:

$$S(x) * F(x) = I(x) \tag{199}$$

The goal is to reconstruct $S(x)$ by "antiblurring" $I(x)$, which is equivalent to convolving $I(x)$ with the convolution inverse of the filter function $F(x)$. In other words, $I(x)$ will be anti-blurred, and the original signal $S(x)$ will be recovered by inverting Eq.(199). To this end, we will break down the signal, the filter, and the blurred signal into wavelet "building blocks", and process the building blocks themselves, instead of working on the entire waveforms. We will then prove that the resulting wavelet equations can be easily inverted, and that the inverse problem is well-posed, with only a mild restriction on the filtering function $F(x)$. We proceed by

- expanding the original $S(x)$ in terms of a truncated HR wavelet expansion of weight λ:

$$S(x) \approx \sum_{n=0}^{N} a_n W_n^{\lambda}(x) \tag{200}$$

where $\{a_n\}_n$ are the Hermite moments of the signal $S(x)$, weighted by λ.

- expanding the filter $F(x)$ in terms of a truncated HR wavelet expansion of weight μ:

$$F(x) \approx \sum_{n=0}^{N} b_n W_n^{\mu}(x) \tag{201}$$

where $\{b_n\}_n$ are the Hermite moments of the filter $F(x)$, weighted by μ.

- expanding the blurred signal $I(x)$ in terms of a truncated HR wavelet expansion of weight ν:

$$I(x) \approx \sum_{n=0}^{N} c_n W_n^{\nu}(x) \tag{202}$$

where $\{c_n\}_n$ are the Hermite moments of the signal $I(x)$, weighted by ν.

Eq.(199) becomes:

$$\left\{\sum_n a_n W_n^{\lambda}(x)\right\} * \left\{\sum_m b_m W_m^{\mu}(x)\right\} = \sum_p c_p W_p^{\nu}(x) \tag{203}$$

From Appendix D, we know that

$$W_n^{\lambda}(x) * W_m^{\mu}(x) = \gamma_{n,m}^{\lambda,\mu}\, W_{n+m}^{\sqrt{\lambda^2+\mu^2}}(x) \tag{204}$$

where

$$\gamma_{n,m}^{\lambda,\mu} = \frac{\lambda^n \mu^m}{(\lambda^2+\mu^2)^{\frac{n+m}{2}}} \sqrt{\frac{(n+m)!}{n!m!}} \tag{205}$$

and thus, writing the subscript $_{n,m}$ as $_{nm}$, and suppressing the superscript $^{\lambda,\mu}$ when confusion does not arise, Eq.(203) becomes:

$$\begin{aligned}
c_0 &= \gamma_{00} a_0 b_0 \\
c_1 &= \gamma_{01} a_0 b_1 + \gamma_{10} a_1 b_0 \\
c_2 &= \gamma_{02} a_0 b_2 + \gamma_{11} a_1 b_1 + \gamma_{20} a_2 b_0 \\
&\vdots \\
c_n &= \sum_{p=0}^{n} \gamma_{p,n-p} a_p b_{n-p} \\
&\vdots
\end{aligned} \tag{206}$$

In matrix notation,

$$\begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \gamma_{00} b_0 & 0 & 0 & \cdots & 0 \\ \gamma_{01} b_1 & \gamma_{10} b_0 & 0 & \cdots & 0 \\ \gamma_{02} b_2 & \gamma_{11} b_1 & \gamma_{20} b_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{0n} b_n & \gamma_{1,n-1} b_{n-1} & \gamma_{2,n-2} b_{n-2} & \cdots & \gamma_{n0} b_0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{207}$$

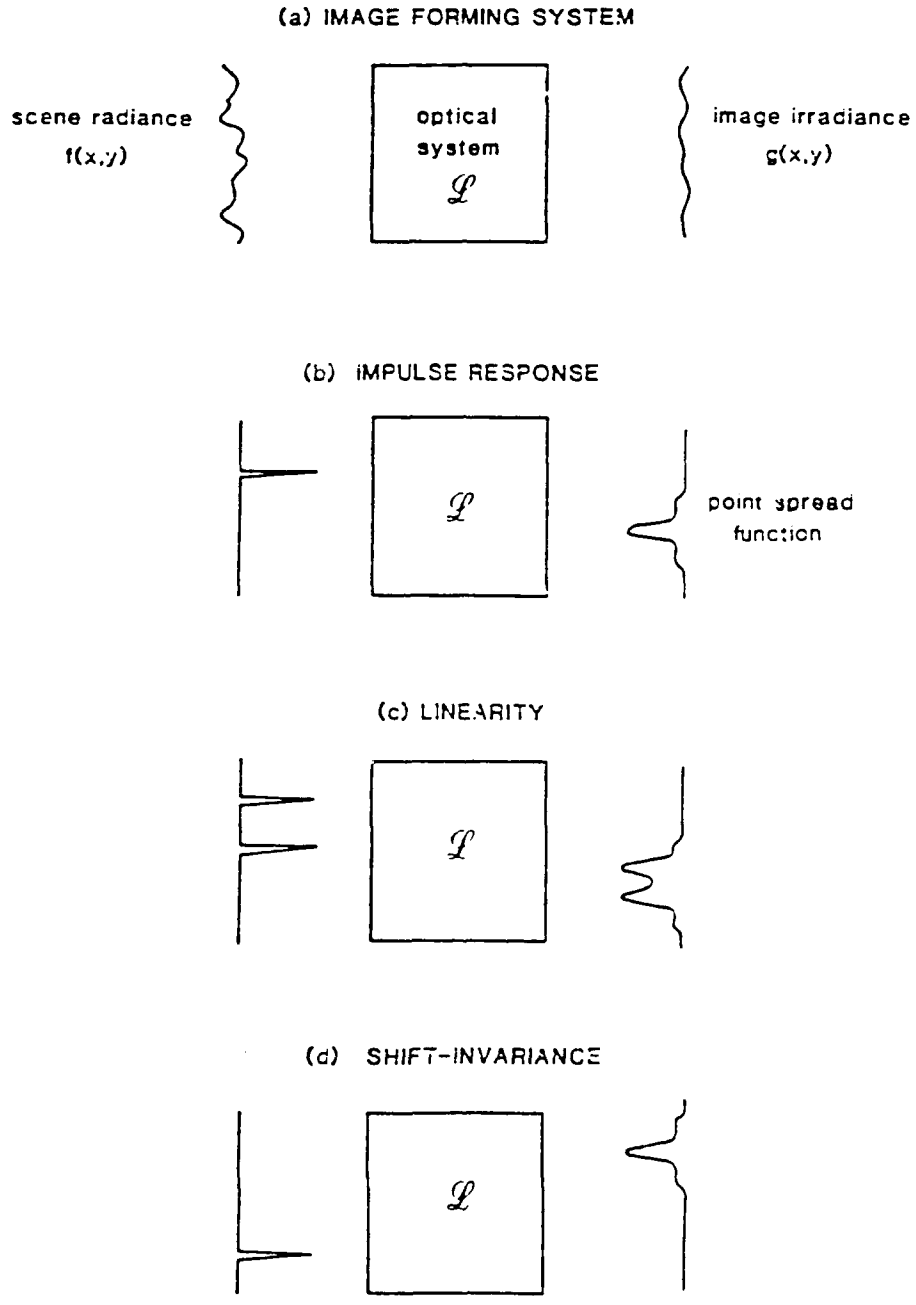
It can be seen that the above matrix is lower triangular, and thus that the determinant is equal to the product of the main diagonal, which is equal to $b_0^n \prod_{k=0}^{n} \gamma_{k0}$. That product is always nonzero, provided that b_0 be non-zero. In other words, the matrix equation (207) can always be inverted, provided that the zeroth moment b_0 of the filter be non-zero. But since we require that the filter conserve the total intensity of the original waveform $S(x)$, b_0 must be equal to 1 (see Appendix F), and Eq.(207) can always be inverted. The solution is obtained by simple Gaussian elimination with the following recursion formula:

$$\begin{aligned} a_0 &= \frac{1}{\gamma_{00} b_0} c_0 \\ a_n &= \frac{1}{\gamma_{n0} b_0} \left\{ c_n - \sum_{p=0}^{n-1} \gamma_{p,n-p} a_p b_{n-p} \right\} \end{aligned} \tag{208}$$

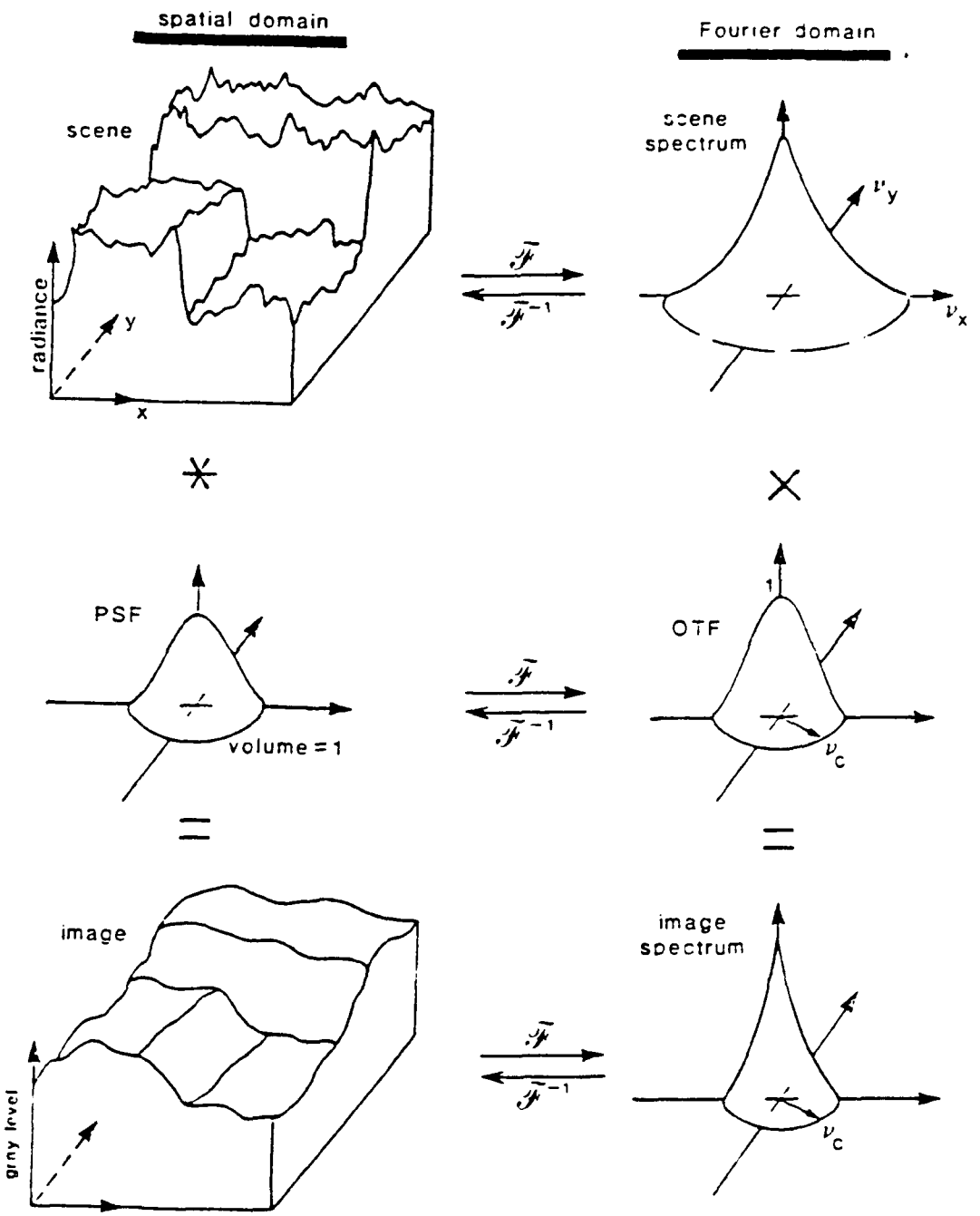
Once the coefficients $\{a_n\}_n$ are determined, the original signal $S(x)$ can be reconstructed by its wavelet expansion:

$$S(x) = \sum_n a_n W_n^{\lambda}(x) \tag{209}$$

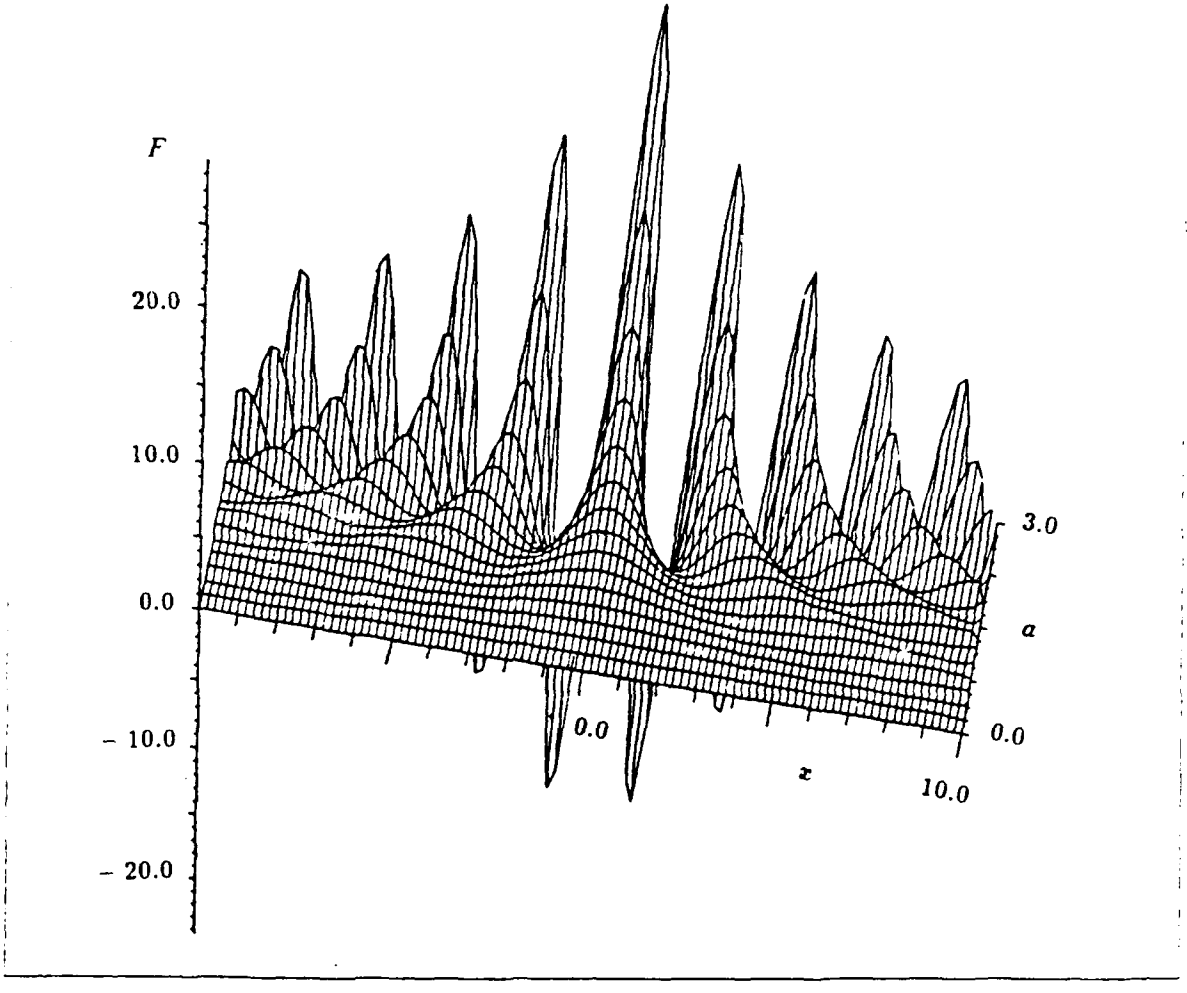
Experiments involving the reconstruction of simulated landscapes and nebulae corrupted with non-gaussian blur are being developed.



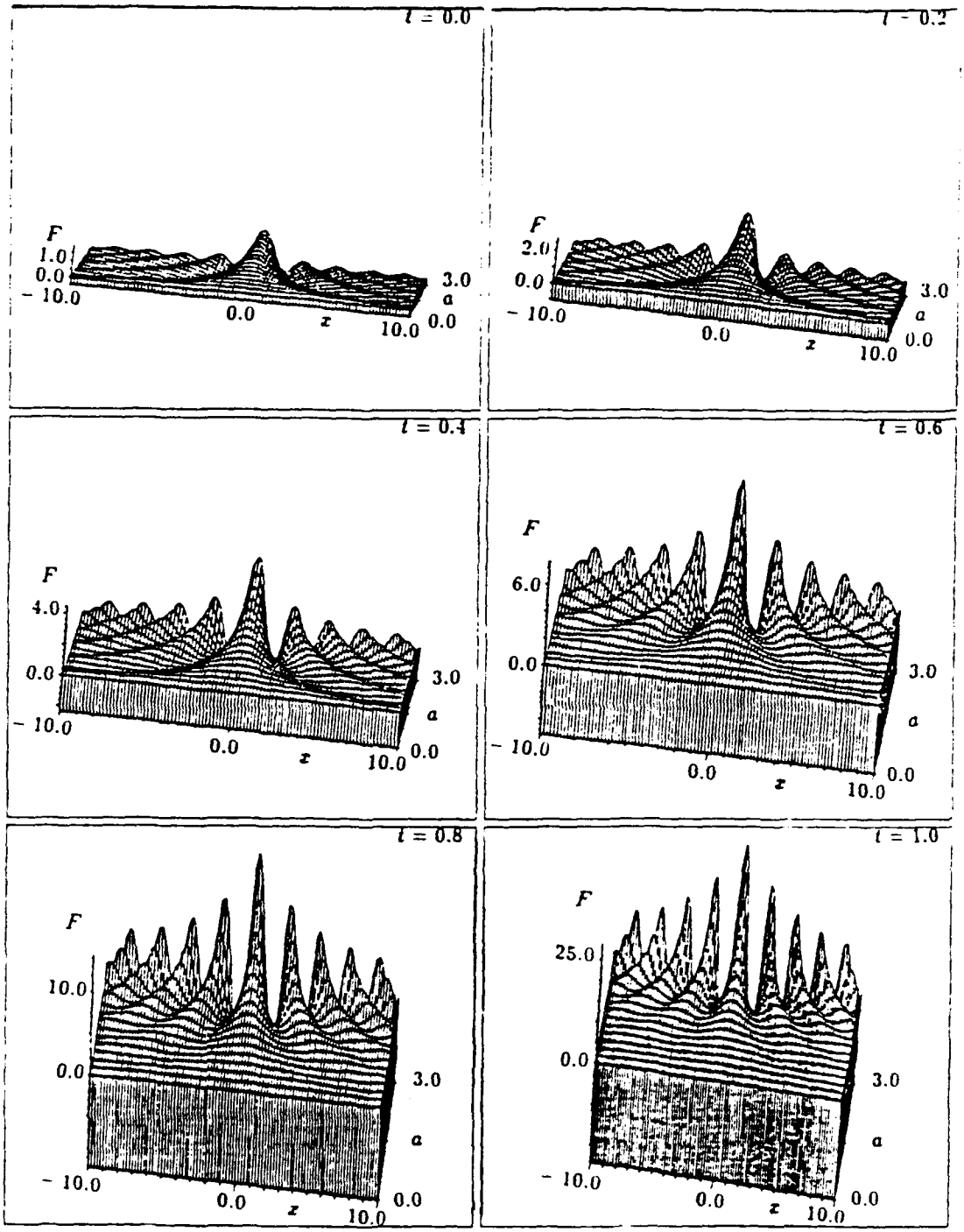
1. Properties of linear and time-invariant (LTI) systems.



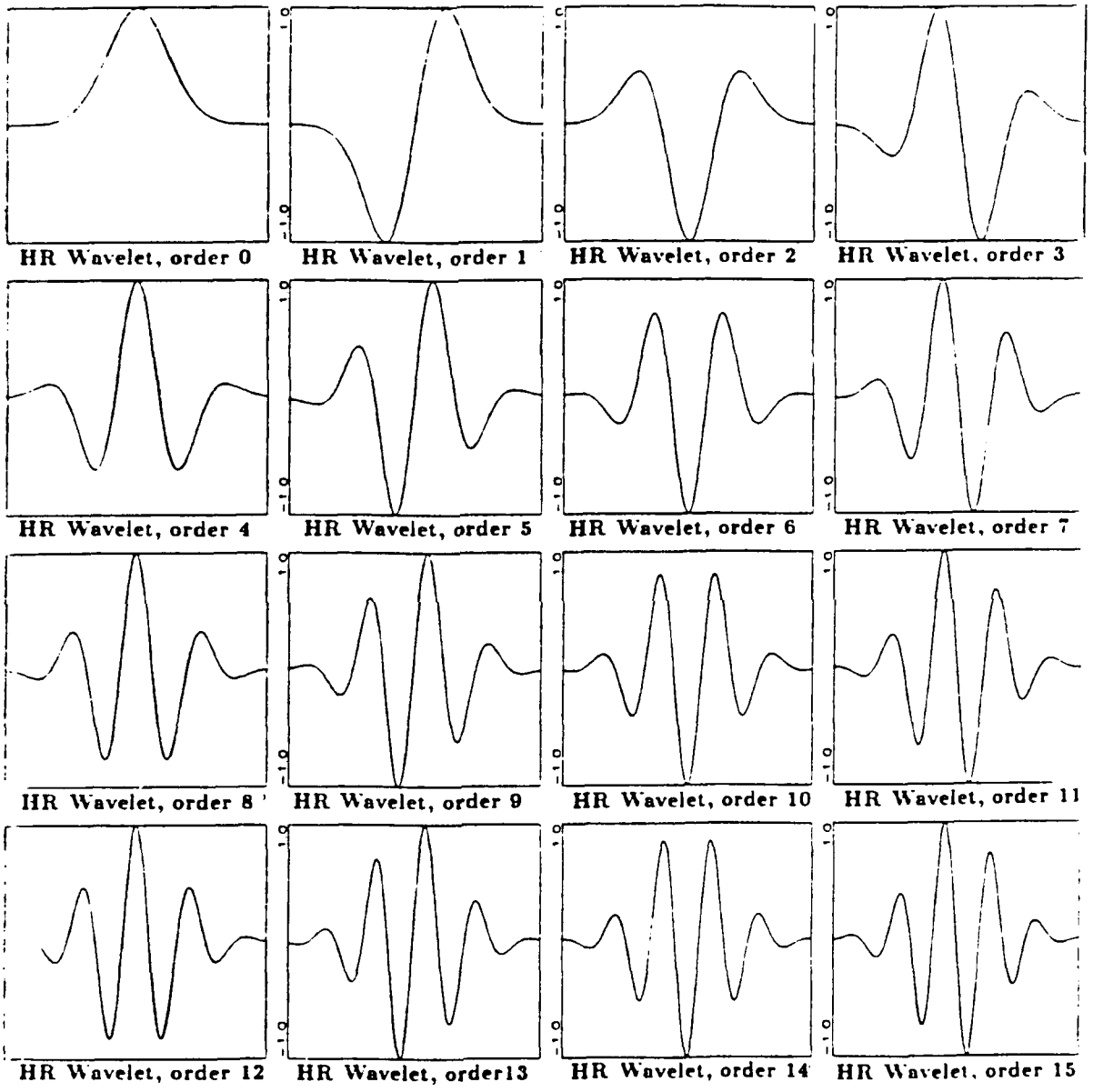
2. Image formation in the spatial and Fourier domains. Filtering by convolution with a smooth point-function is a "smoothing" operation.



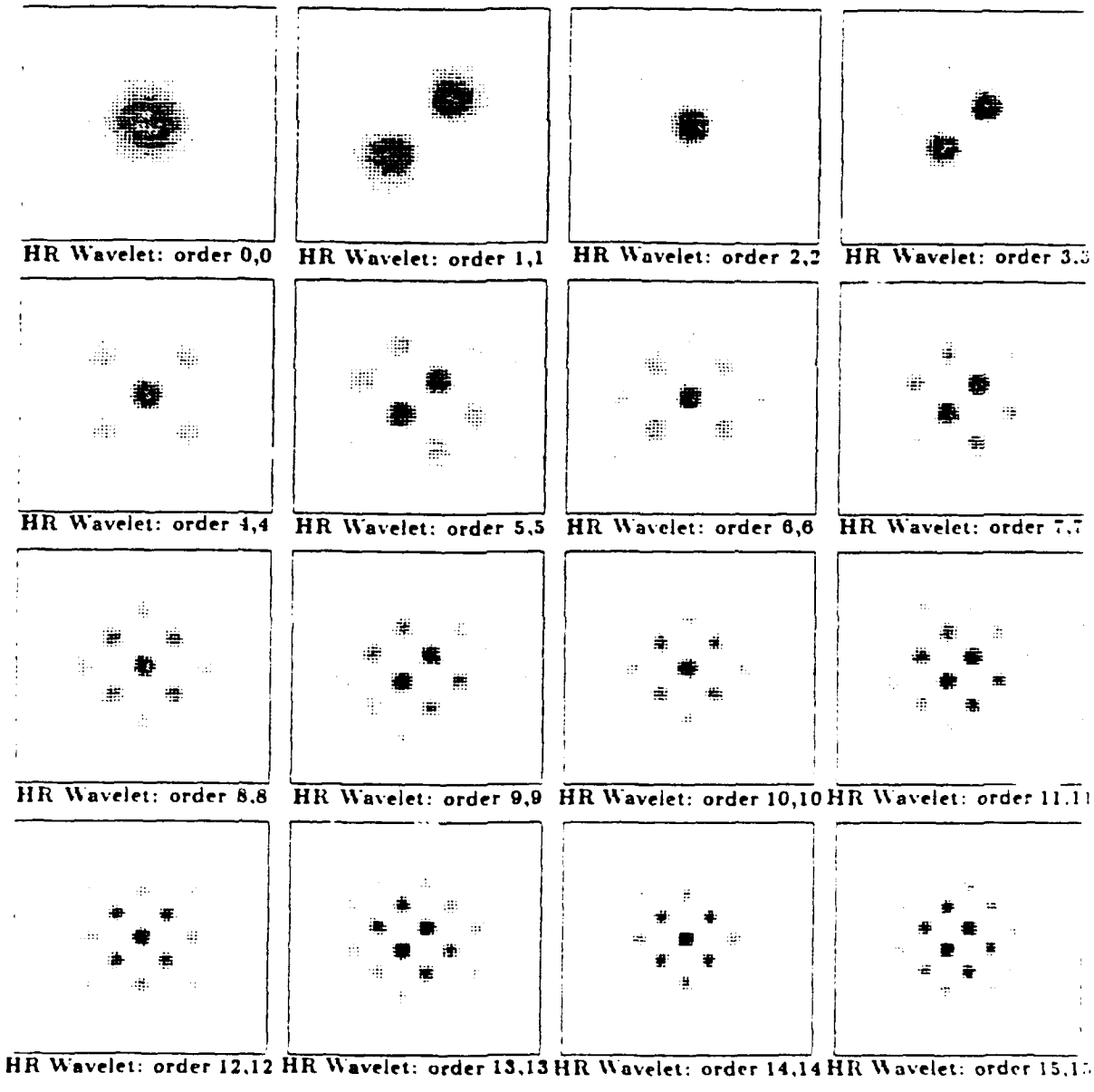
3. Approximations of the convolution inverse of a gaussian filter of unit width, as given by Fourier methods. The limits of the Fourier integral are cut-off at $\pm a$: $F(a,x,1) = \frac{1}{2\pi}\int_{-a}^{+a} e^{-ikx+\frac{k^2}{2}}\,dk$. The vertical F axis' endpoints are the extrema of $F(3.0,x,1)$.



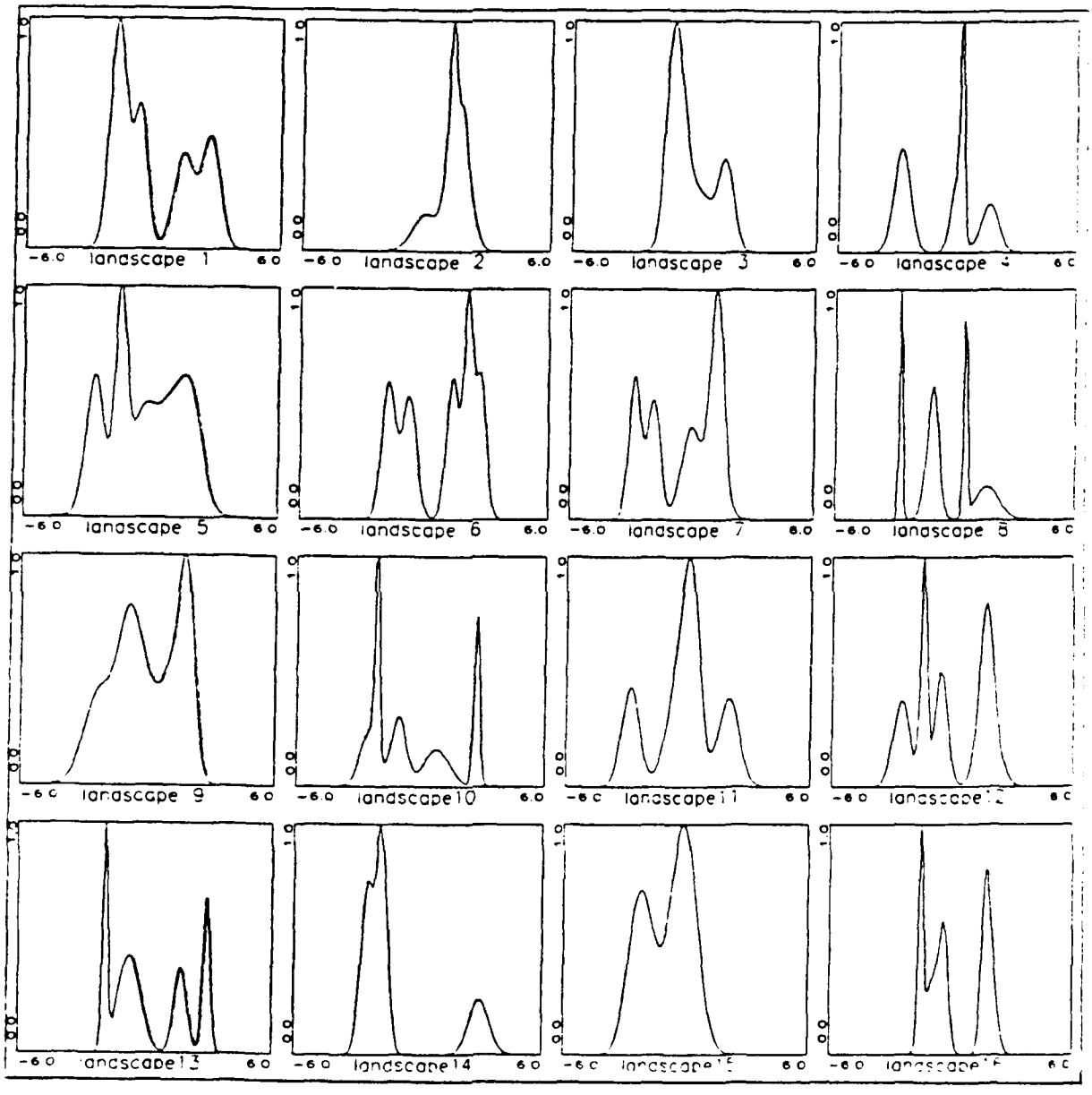
4. Approximations of the convolution inverse of a gaussian filter of increasing width t, as given by Fourier methods. The limits of the Fourier integral are cut-off at $\pm a$: $F(a, x, \frac{t}{2}) = \frac{1}{2\pi} \int_{-a}^{+a} e^{-ikx + \frac{k^2 t}{2}}\, dk$. The skirt in each figure ends at the minimum of $F(3.0, x, \frac{t}{2})$, while the upper endpoint of the vertical F axis ends at its maximum.



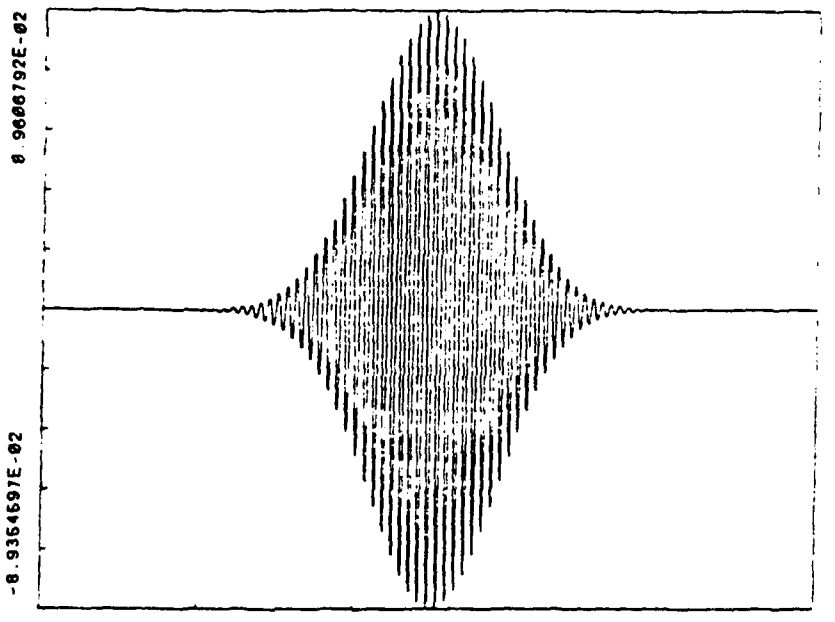
5. Elements of the sequence of 1D HR wavelets of weight $\lambda = 2$, evaluated in the range [-6,6]. The HR wavelets are rescaled to unit height.



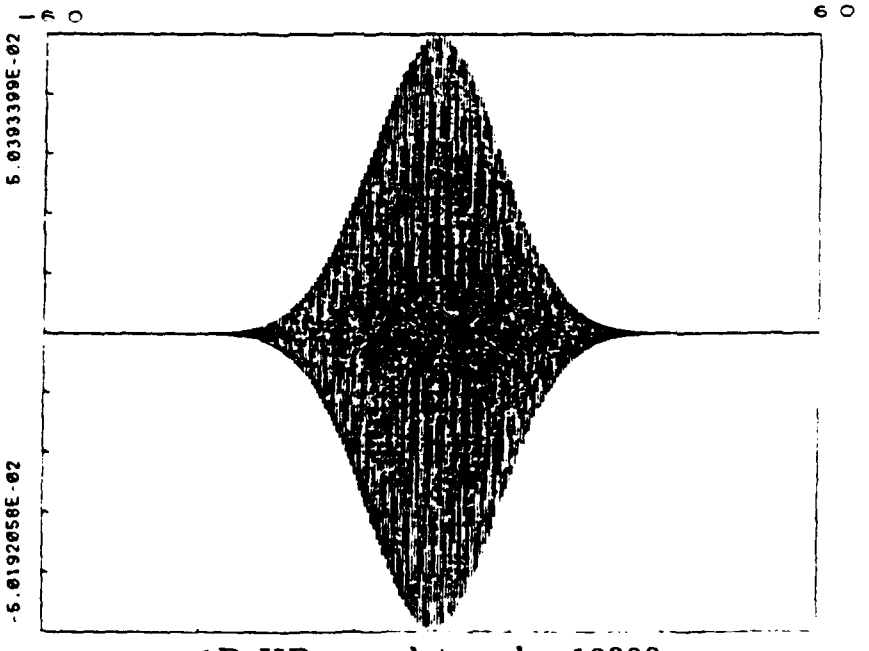
6. Elements of the sequence of 2D HR wavelets of weight $\lambda = \mu = 2$, evaluated in the range [-6,6]×[-6,6], with the ordinate in the range [0,1] exhibited in grey scales (i.e., negative intensities are *not* plotted). The HR wavelets are rescaled to unit maximum intensity.



7. First 16 elements of a family of 80 artificial "landscapes". The landscapes were obtained by summing five gaussians whose widths and offsets were obtained by a random-number generator in an appropriate range. The landscapes are subsequently rescaled to unit height.

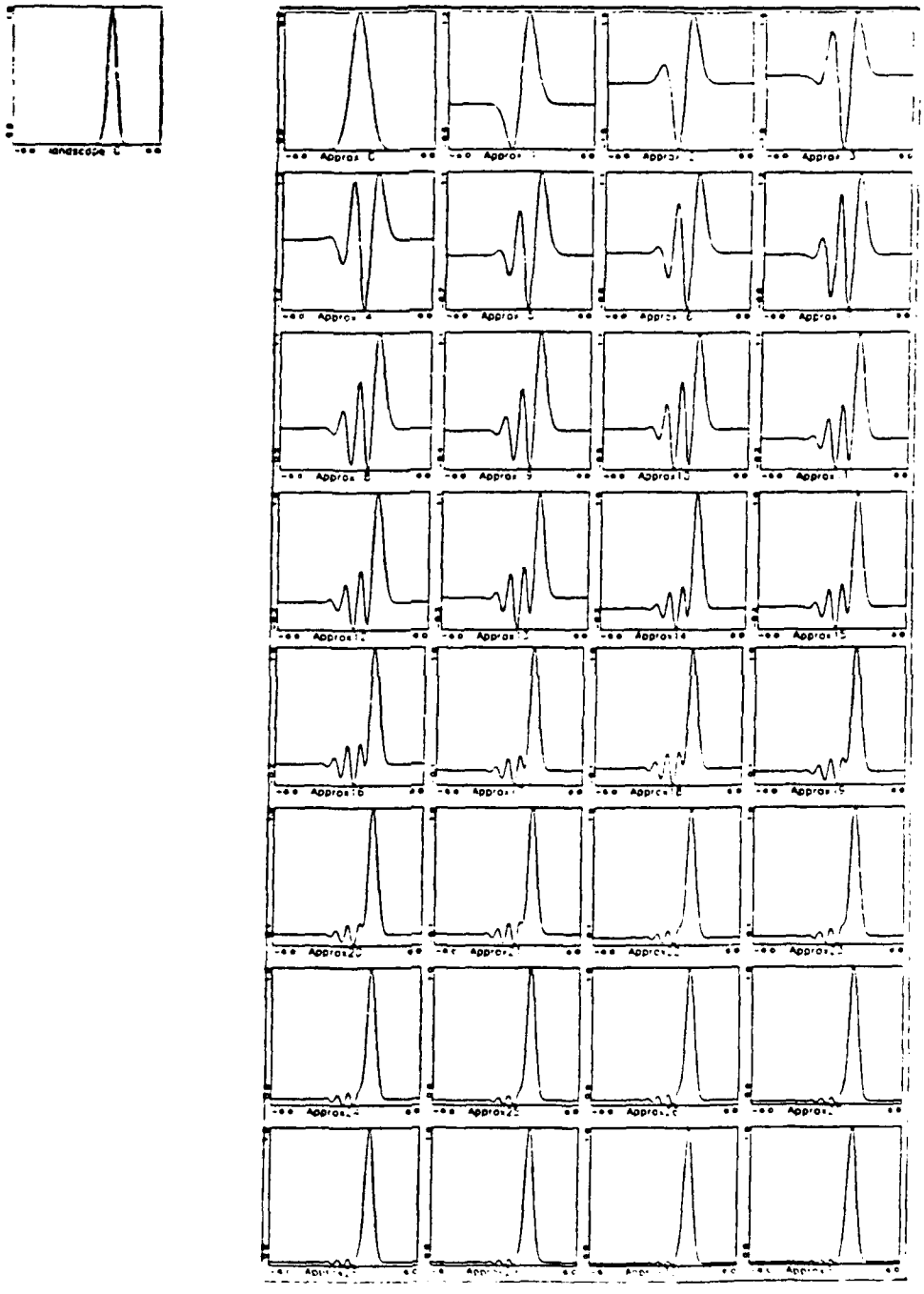


1D HR wavelet, order 1000

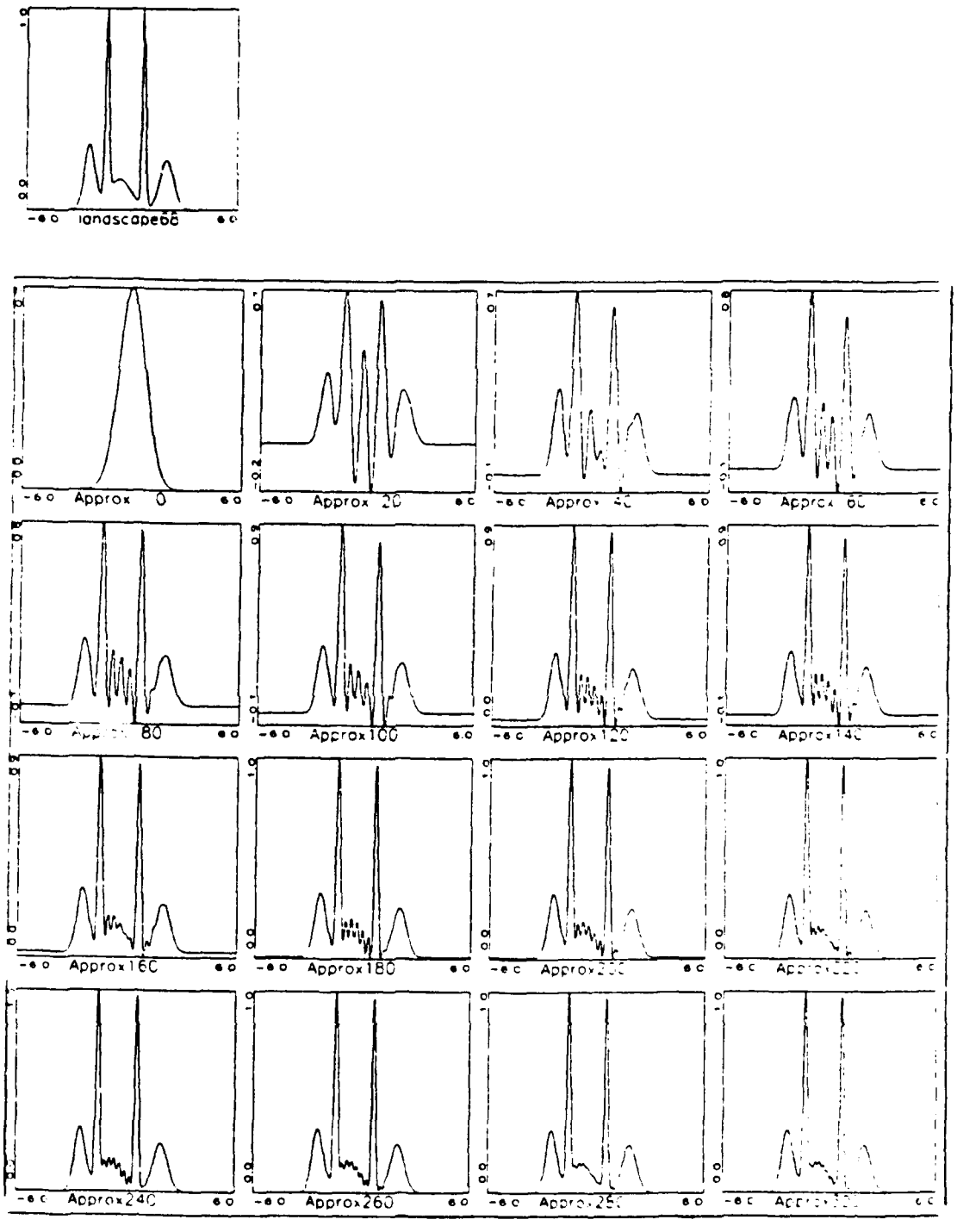


1D HR wavelet, order 10000

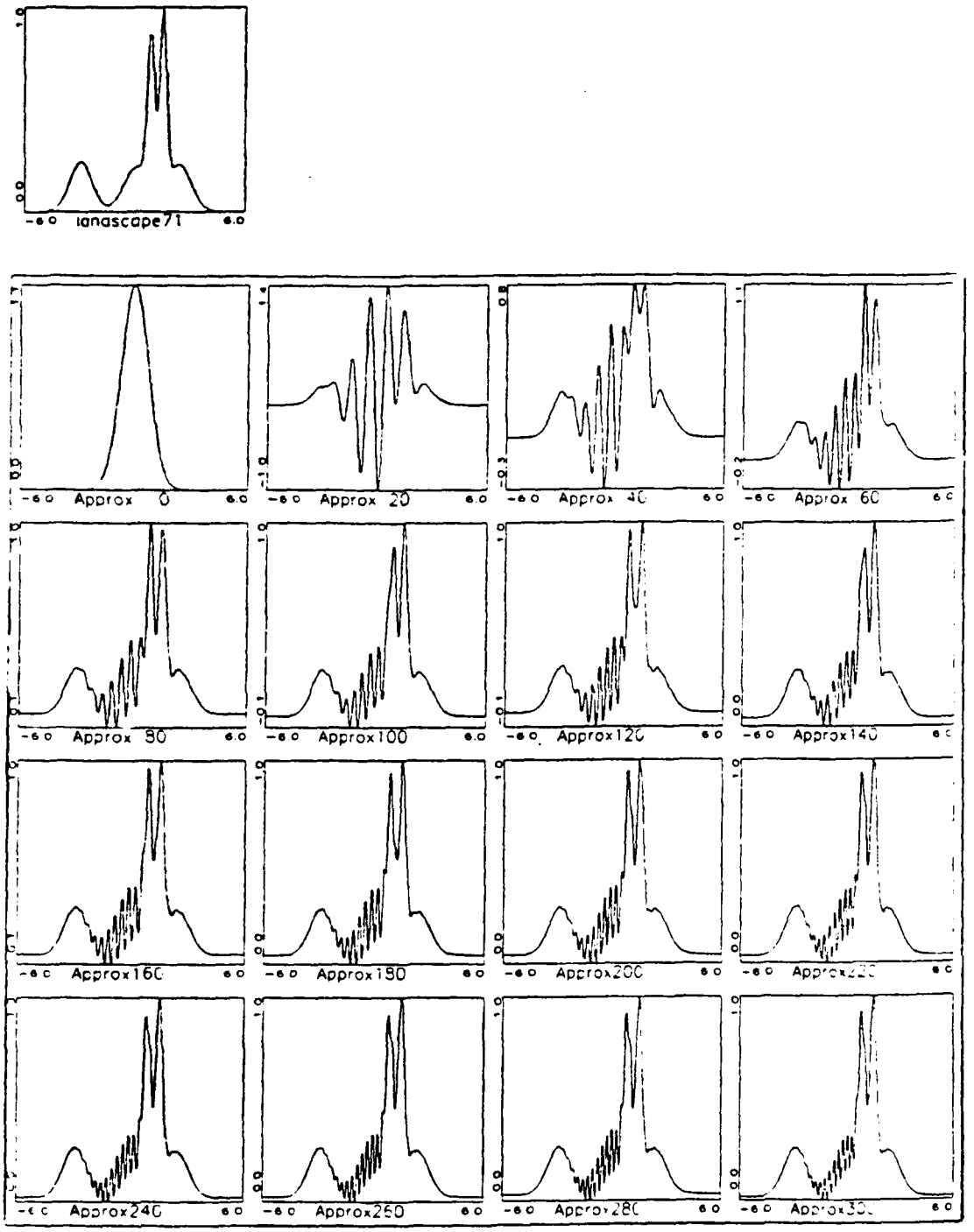
8. 1D HR wavelets of order 1000 and 10000 respectively, and of weight $\lambda = 1$. The renormalization of Hermite polynomials to Hd polynomials allows the numerical evaluation of HR wavelets of high order by eliminating the need for any evaluation of factorial products.



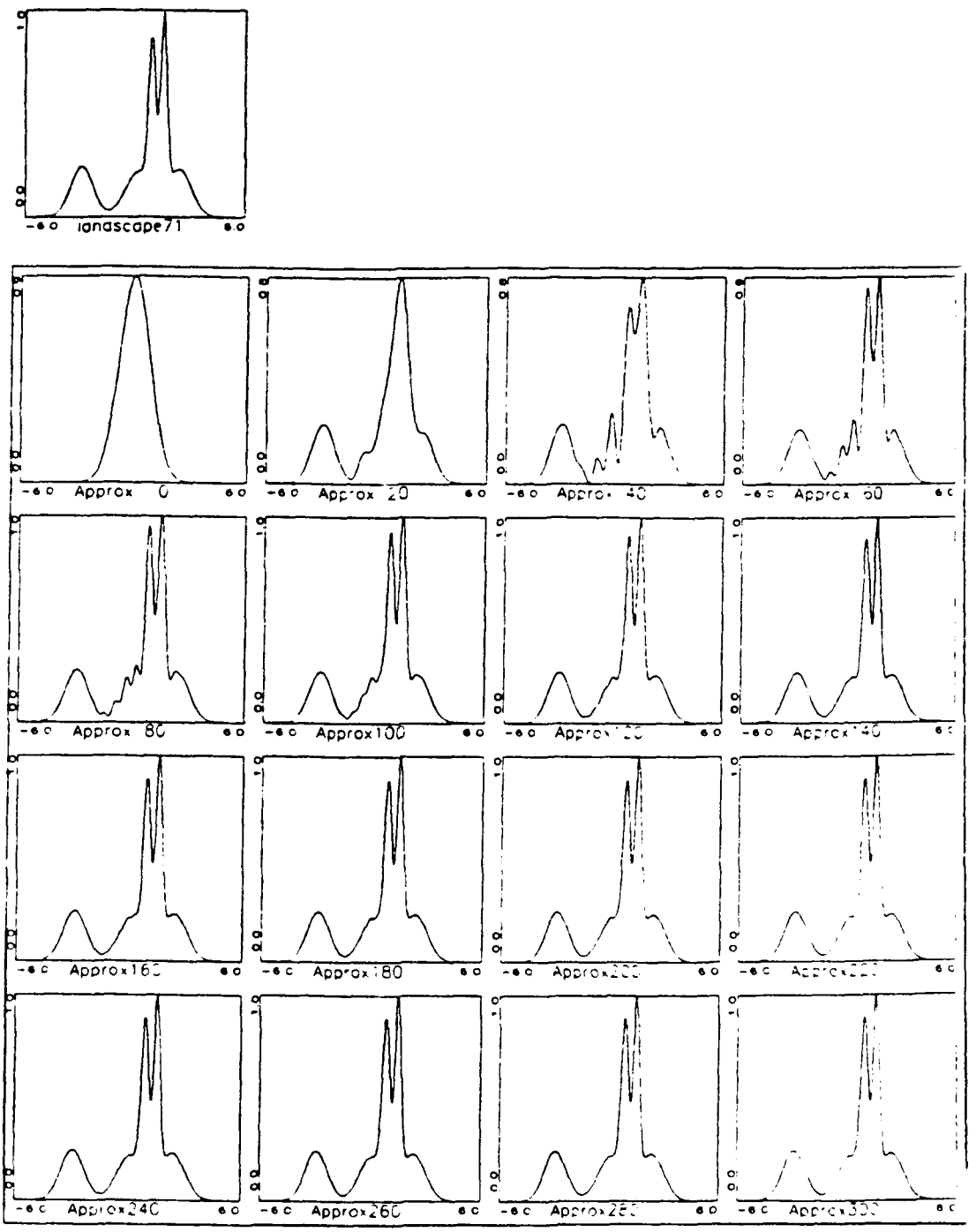
9. A Gaussian of width 0.5, offset at $x = 2$, is approximated by an HR wavelet expansion of width $\lambda = 1$. 32 partial sums are plotted.



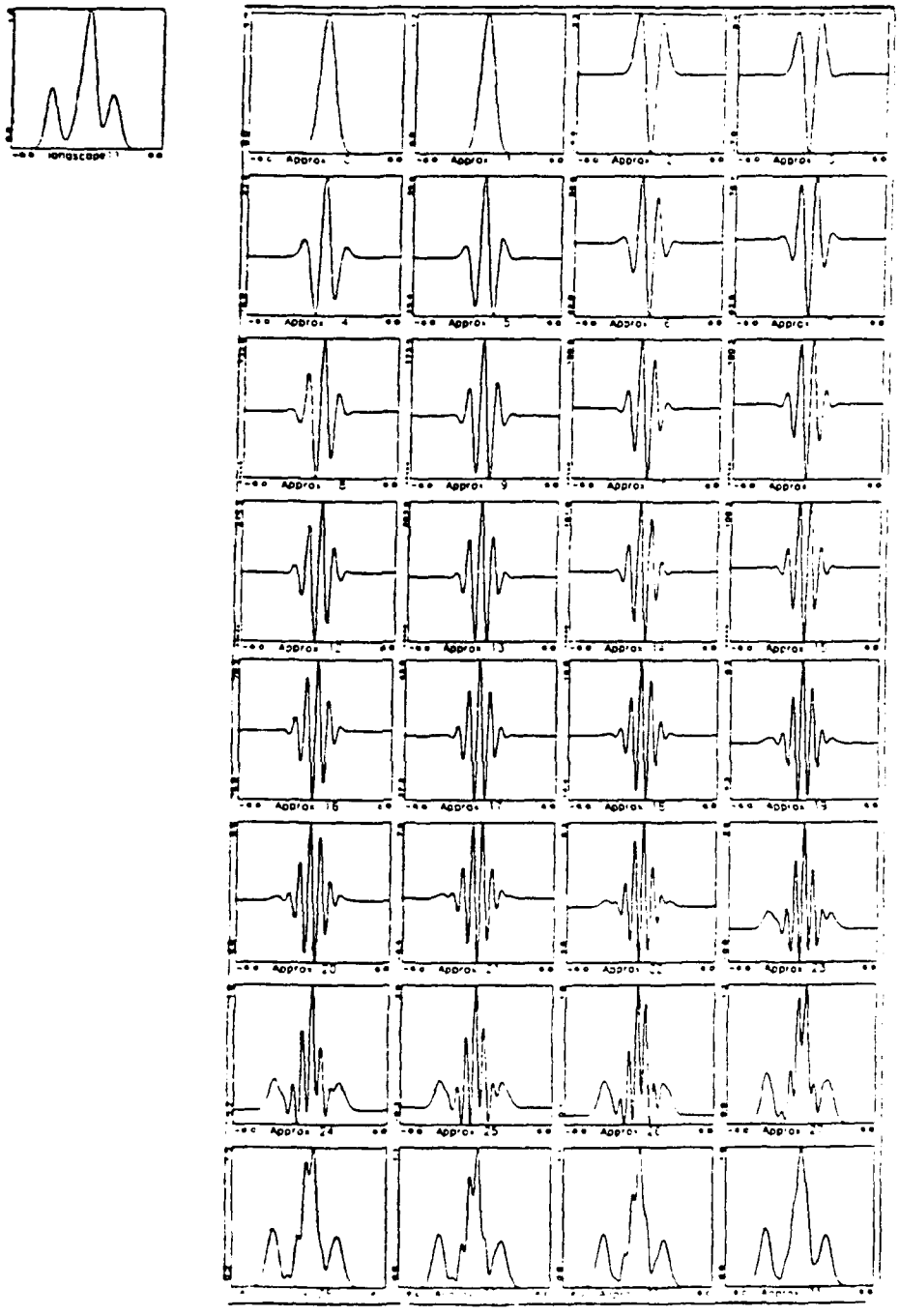
10. The "futuristic" landscape 68 is approximated by an HR wavelet expansion of width $\lambda = 1$. The partial sums are plotted in increments of 20. The capturing of very localized features (i.e. high frequency content) is slower than the capturing of global behavior (i.e. low frequency content features).



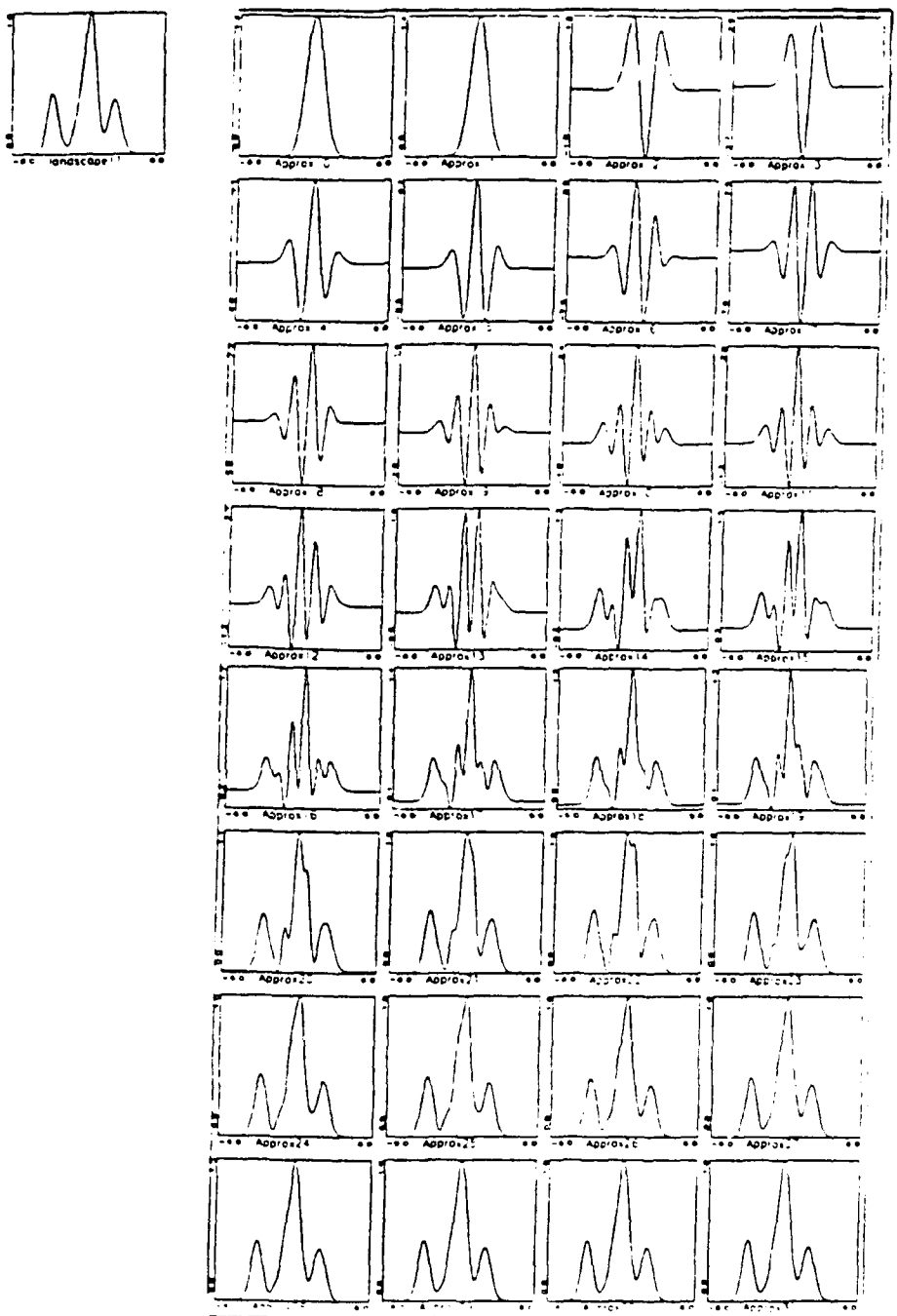
11. The "futuristic"-looking landscape 71 is approximated by an HR wavelet expansion of width $\lambda = 1$. The partial sums are plotted in increments of 20. The wavelet expansion experiences difficulty in approximating one specific feature of the landscape. This figure is to be compared with the next figure.



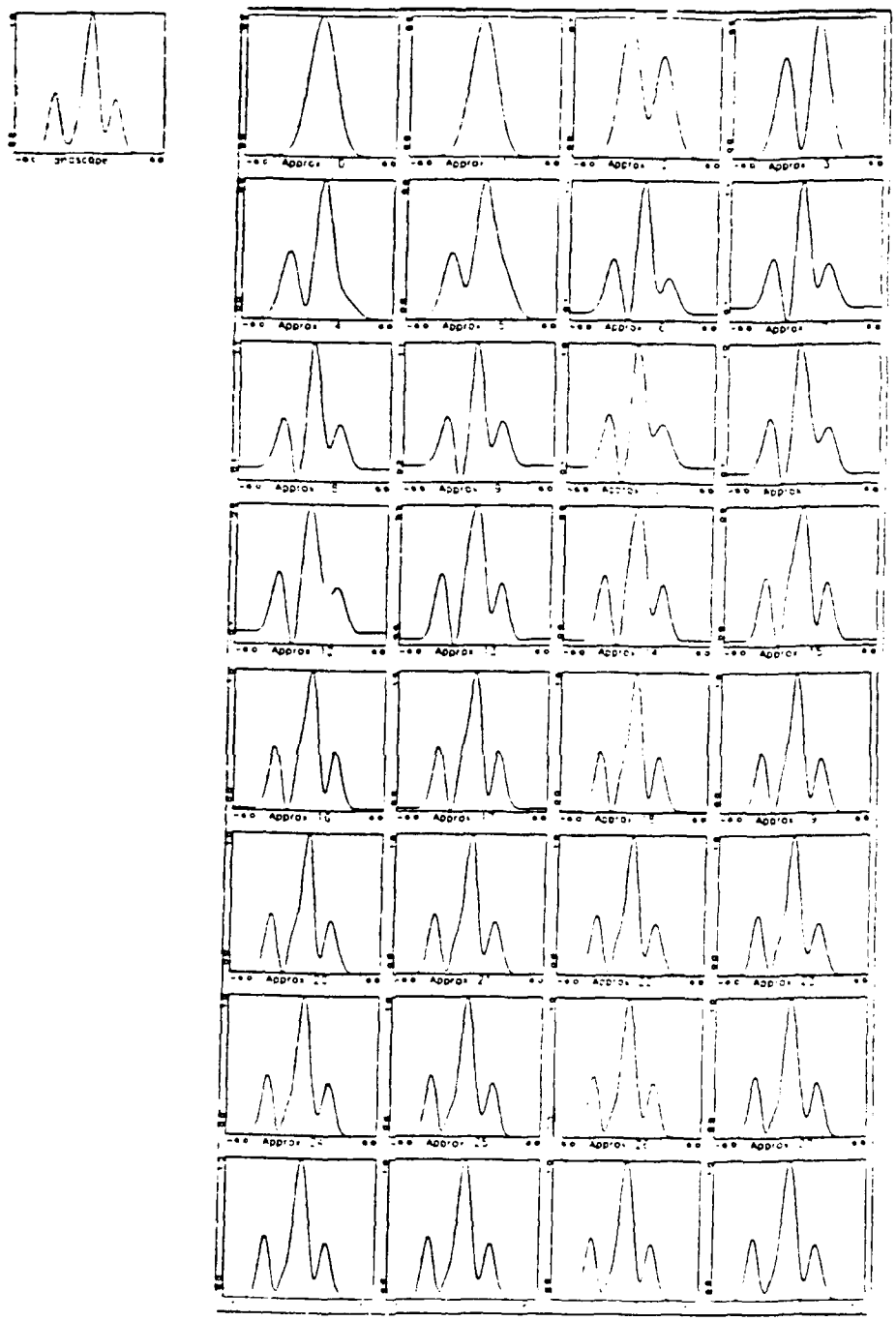
12. The "futuristic"-looking landscape 71 is approximated by an HR wavelet expansion of width $\lambda = 1.15$. The partial sums are plotted in increments of 20. This wavelet expansion, as opposed to the previous expansion for which $\lambda = 1$, now quickly captures all features of the landscape. The text explains why this happens.



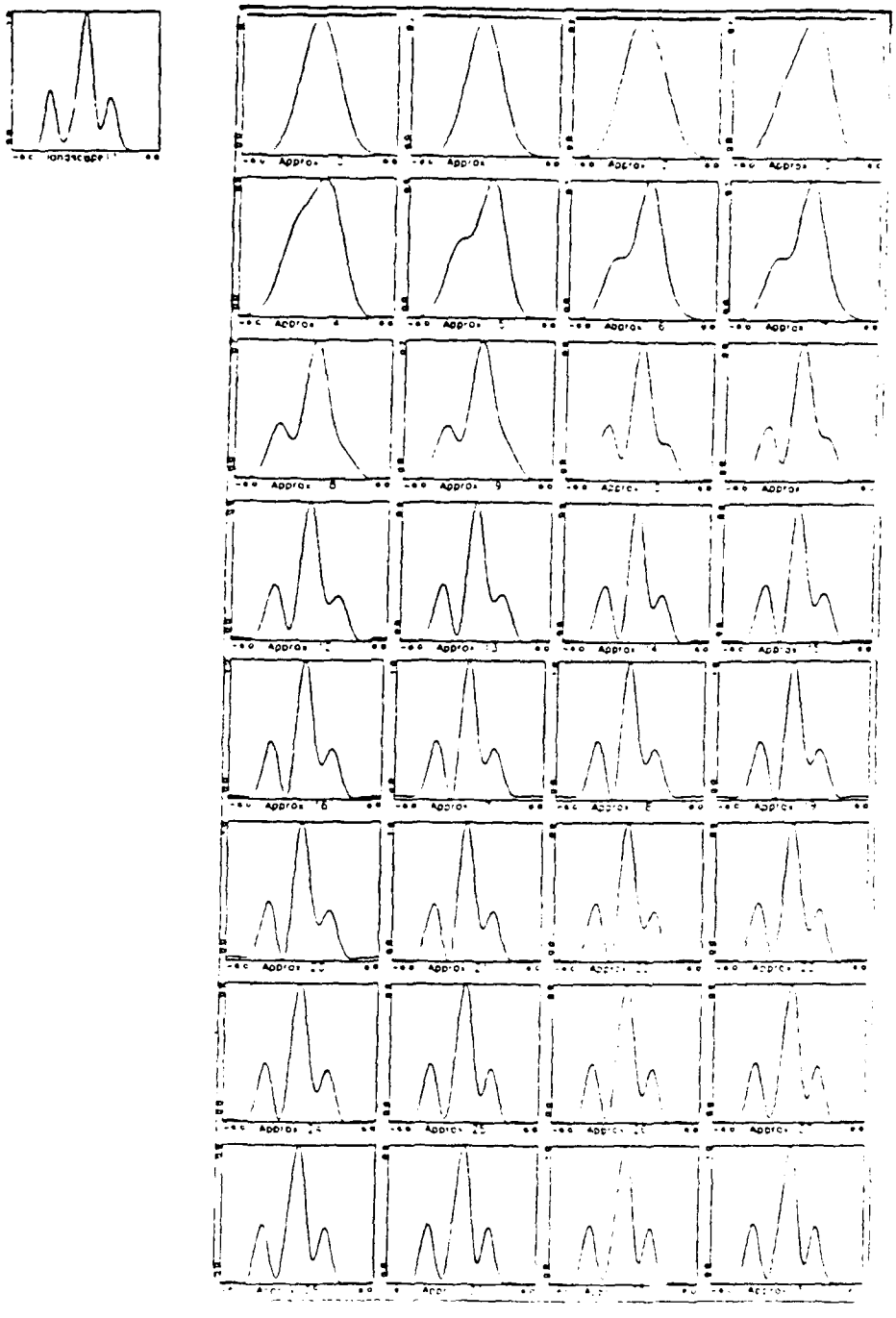
13. The "realistic"-looking landscape 11 is approximated by an HR wavelet expansion of width $\lambda = 0.82$. The partial sums are plotted in increments of 1. The first 21 partial sums are entirely undifferentiated, and look exactly like the sequence of underlying HR wavelets. The approximation starts taking shape with the 23^{rd} partial sum.



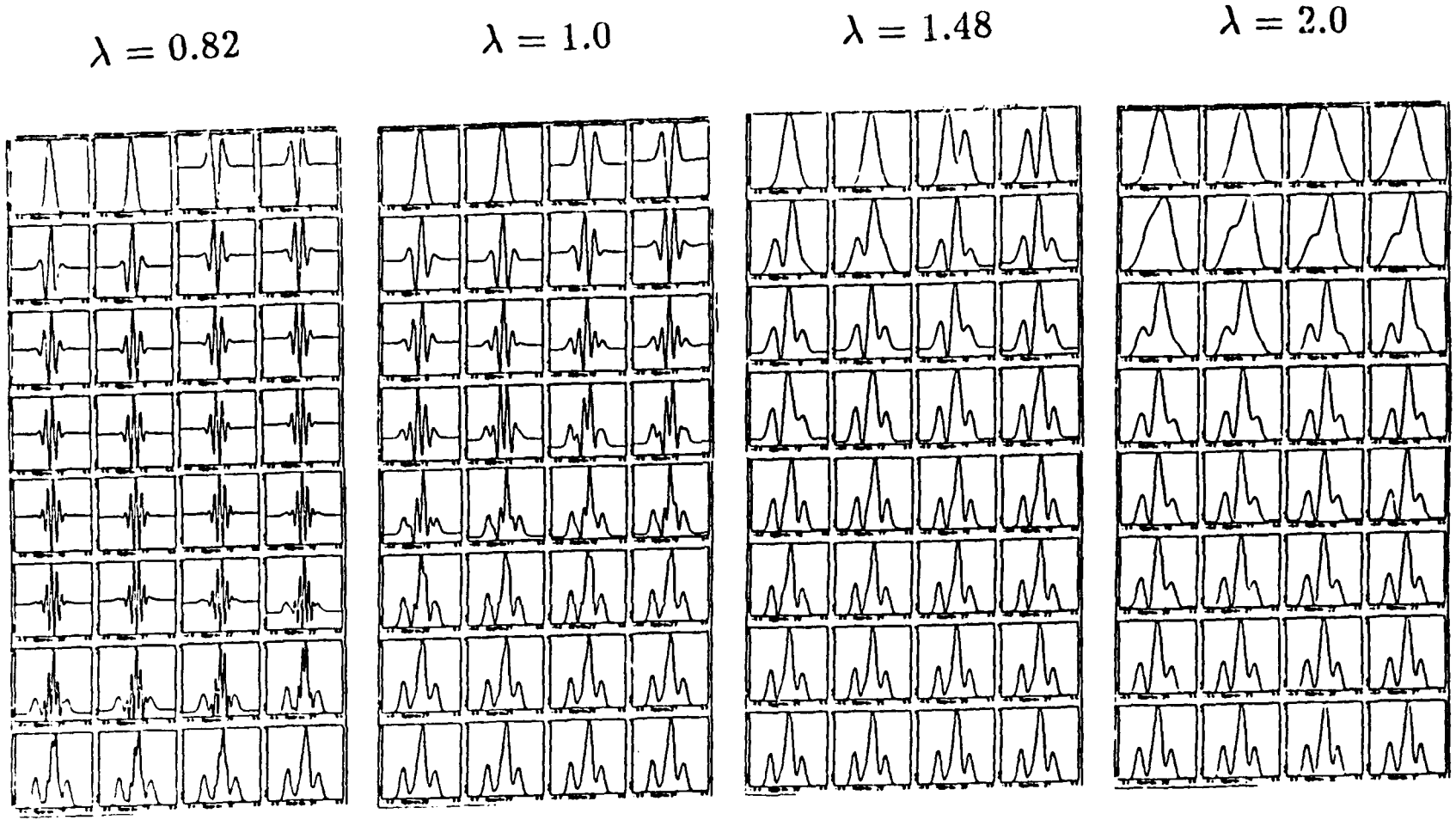
14. The "realistic"-looking landscape 11 is approximated by an HR wavelet expansion of width $\lambda = 1.0$. The partial sums are plotted in increments of 1. The first 9 partial sums are entirely undifferentiated, and look exactly like the sequence of underlying HR wavelets. The approximation starts taking shape with the 13^{th} partial sum.



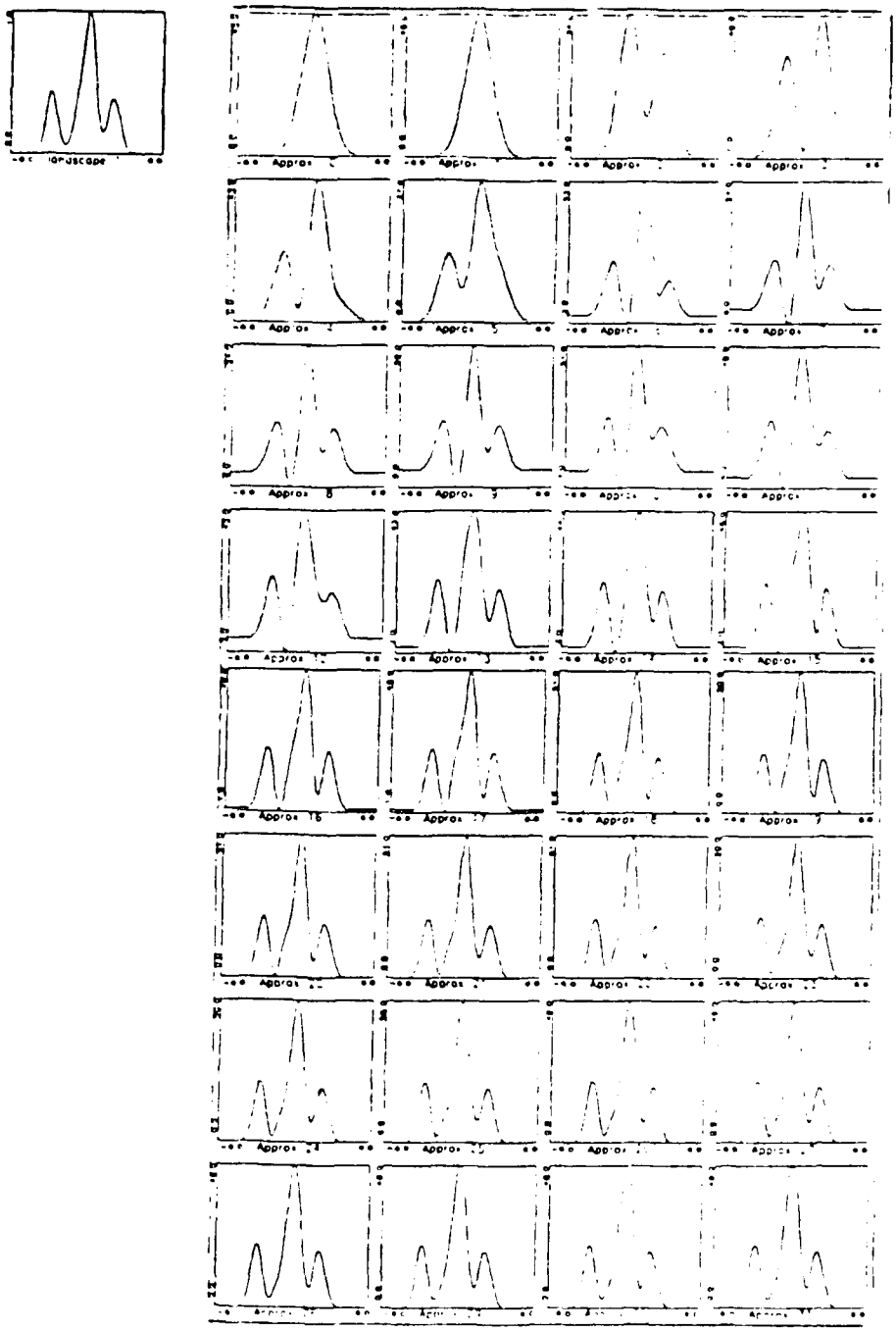
15. The "realistic"-looking landscape 11 is approximated by an HR wavelet expansion of width $\lambda = 1.48$. The partial sums are plotted in increments of 1. The features of landscape 11 are generated with very few partial sums. It takes only 7 partial sums to recognize the landscape. The weight $\lambda = 1.48$ of the wavelet expansion is optimal for this landscape.



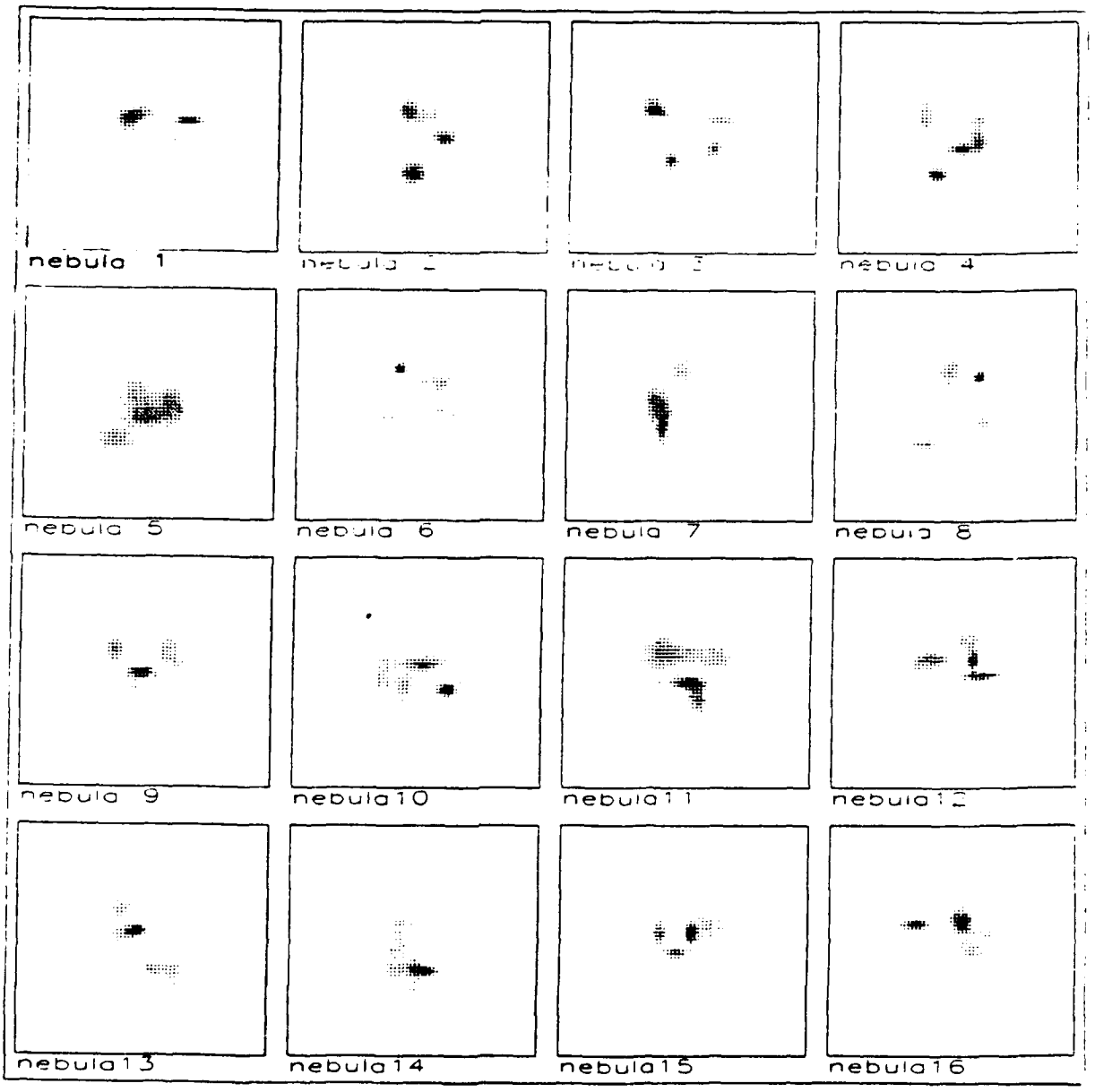
16. The "realistic"-looking landscape 11 is approximated by an HR wavelet expansion of width $\lambda = 2.0$. The partial sums are plotted in increments of 1. λ is once again in a suboptimal range, as it takes at least twelve partial sums to recognize the landscape.



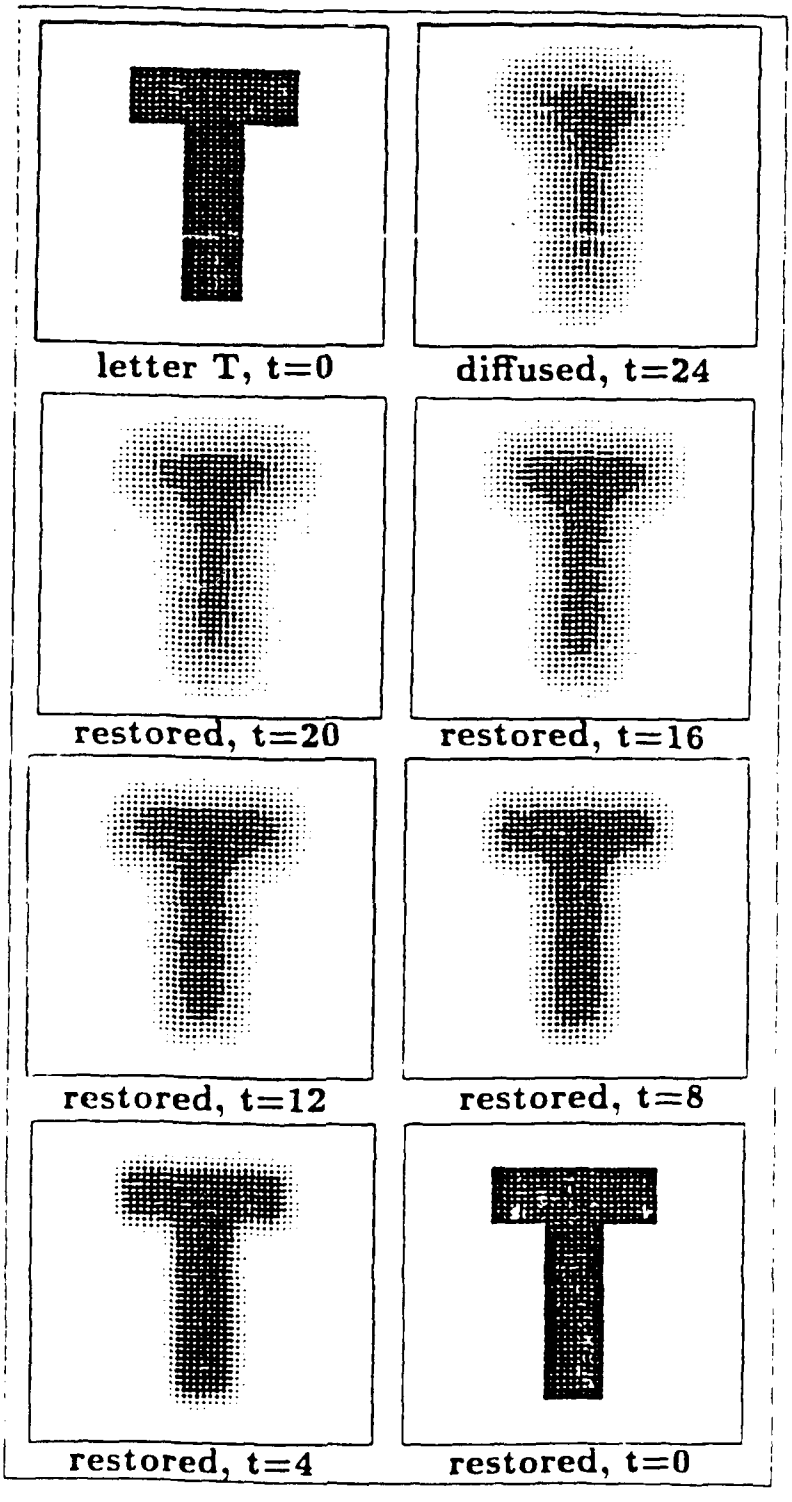
17. The last four figures are summarized. $\lambda = 1.48$ is the optimal value for the IIR wavelet expansion of landscape 11, since the rate of convergence to the landscape is maximum. $\lambda = 0.82$, 1.0, and 2.0 are suboptimal values.



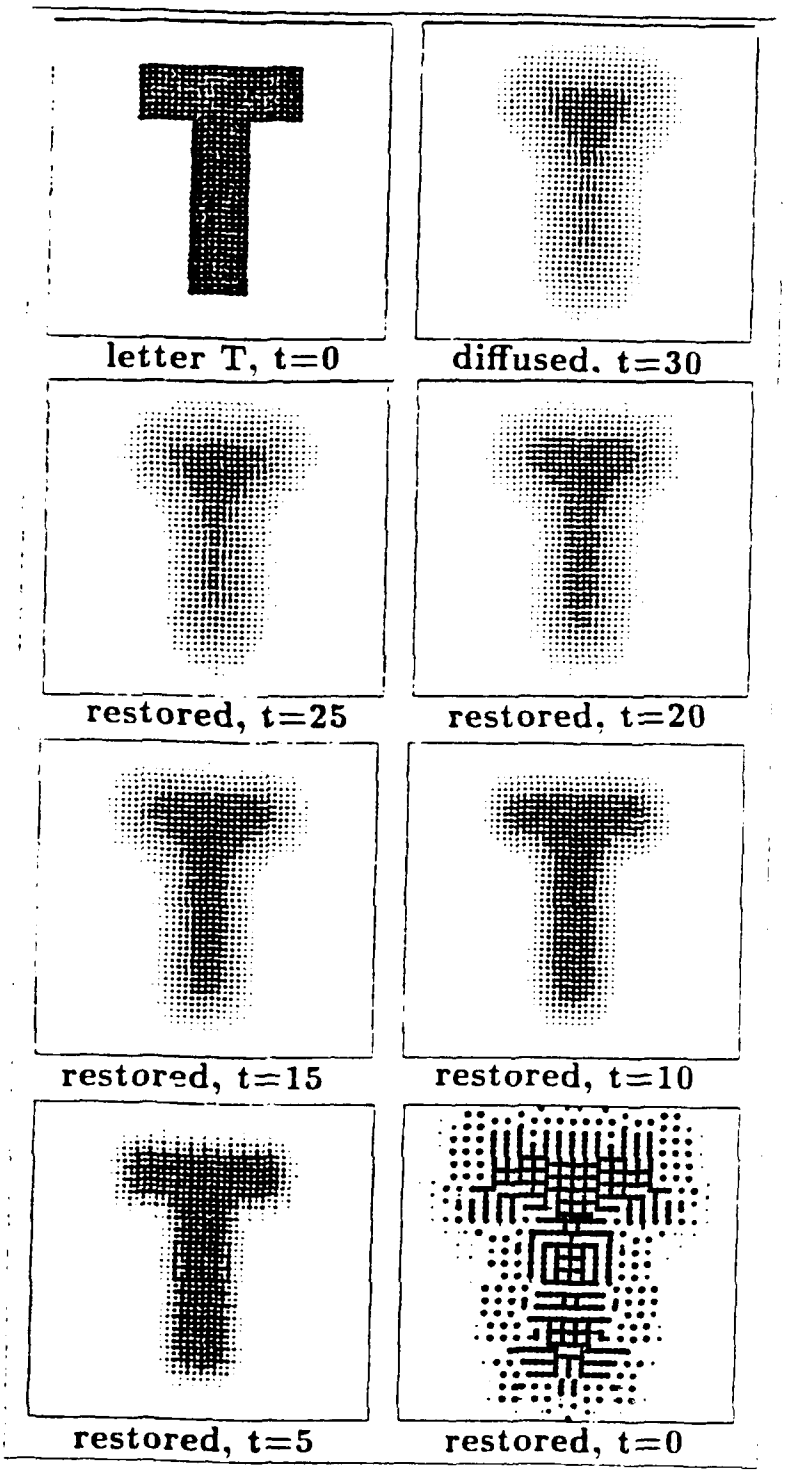
18. The HR moments for landscape 11 were quantized to a range of $256 = 2^8$ values. In other words, the values of the HR moments were slightly modified to fall into 256 equally spaced bins in the range of the HR moments. The weight λ of the expansion is 1.48 (optimal). The rate of convergence of the quantized wavelet expansion is only slightly modified. The HR wavelet expansion is *robust* with respect to small errors on the HR moments.



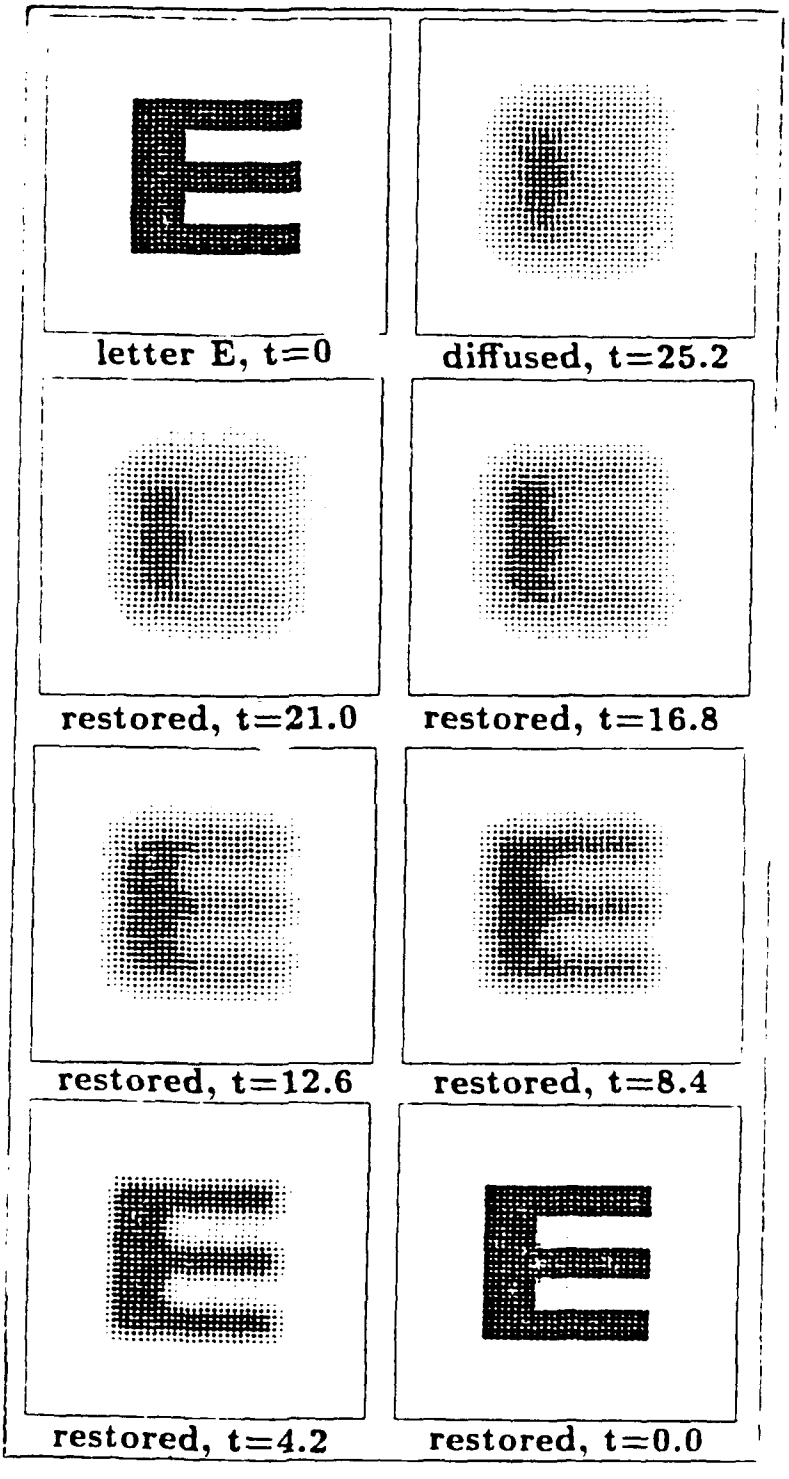
19. First 16 elements of a family of 80 artificial "nebulae". The nebulae were obtained by summing five 2D gaussians whose widths and offsets were obtained by a random-number generator in an appropriate range. The nebulae are subsequently rescaled to unit maximum intensity.



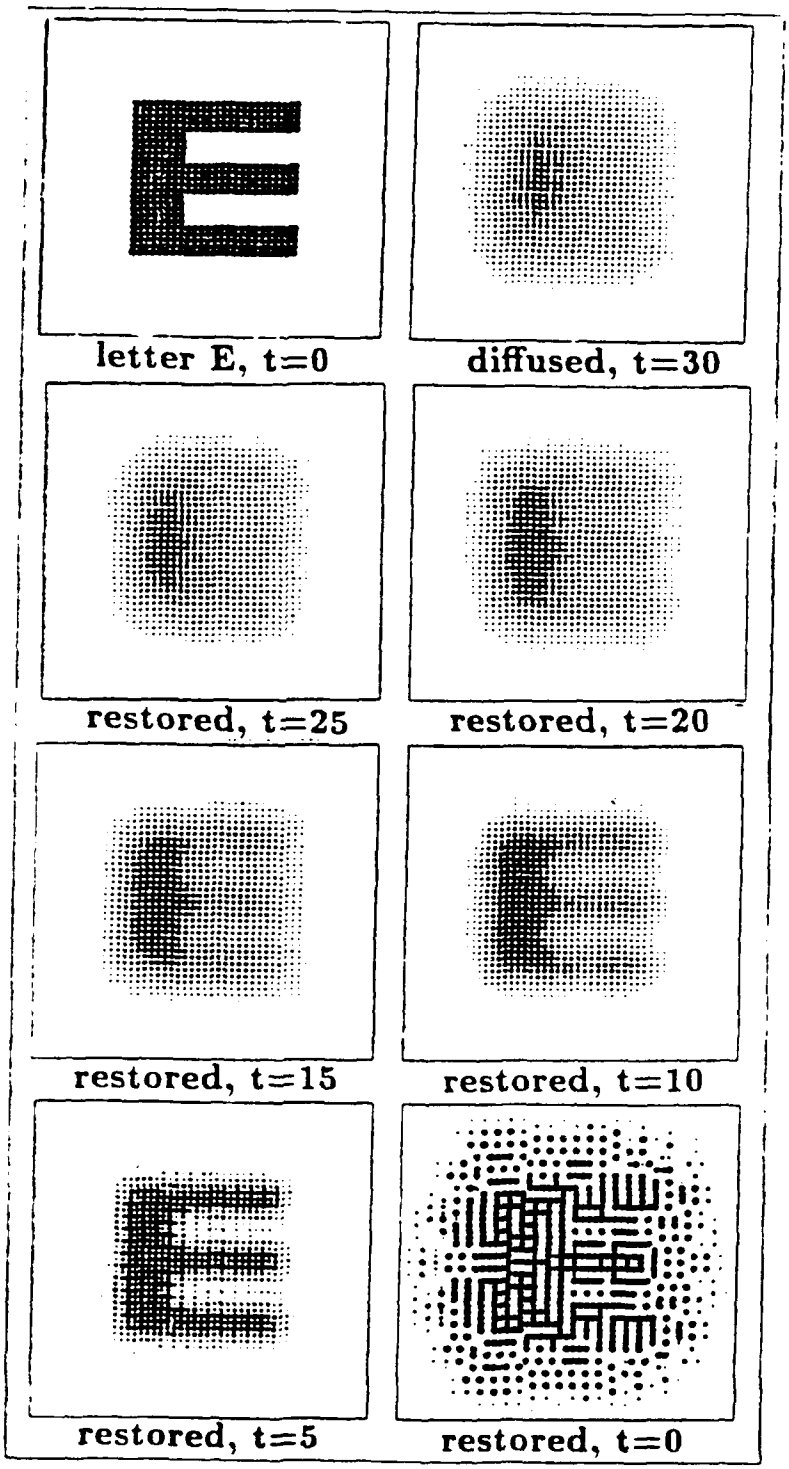
20. Letter "T" (for Tina), which has been diffused for a duration of time $t = 24$, corresponding to 12×10^4 numerical iterations of the discretized diffusion equation (with unit diffusivity), and subsequently reconstructed back to $t = 0$ (with intermediary results) by the same number of iterations of the discretized antidiffusion equation (with unit diffusivity). The reconstruction is successful. The text gives the correct scale of time t.



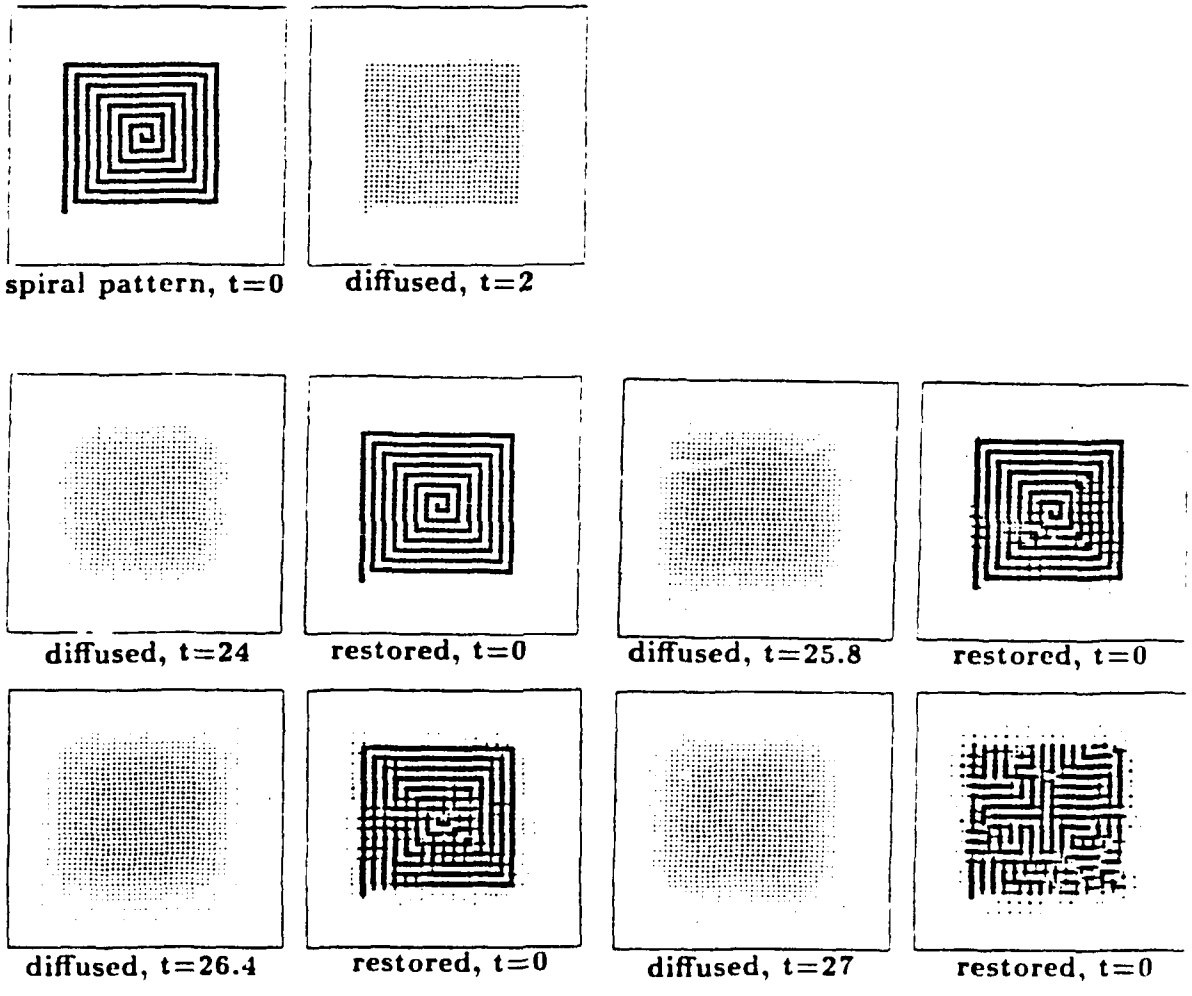
21. Letter "T", which has been diffused for a duration of time $t = 30$ (15×10^4 iterations of the discretized diffusion equation), and subsequently reconstructed back to $t = 0$ (with intermediary results) by iterating the discretized antidiffusion equation (15×10^4 iterations). The $t = 0$ reconstruction is not successful.



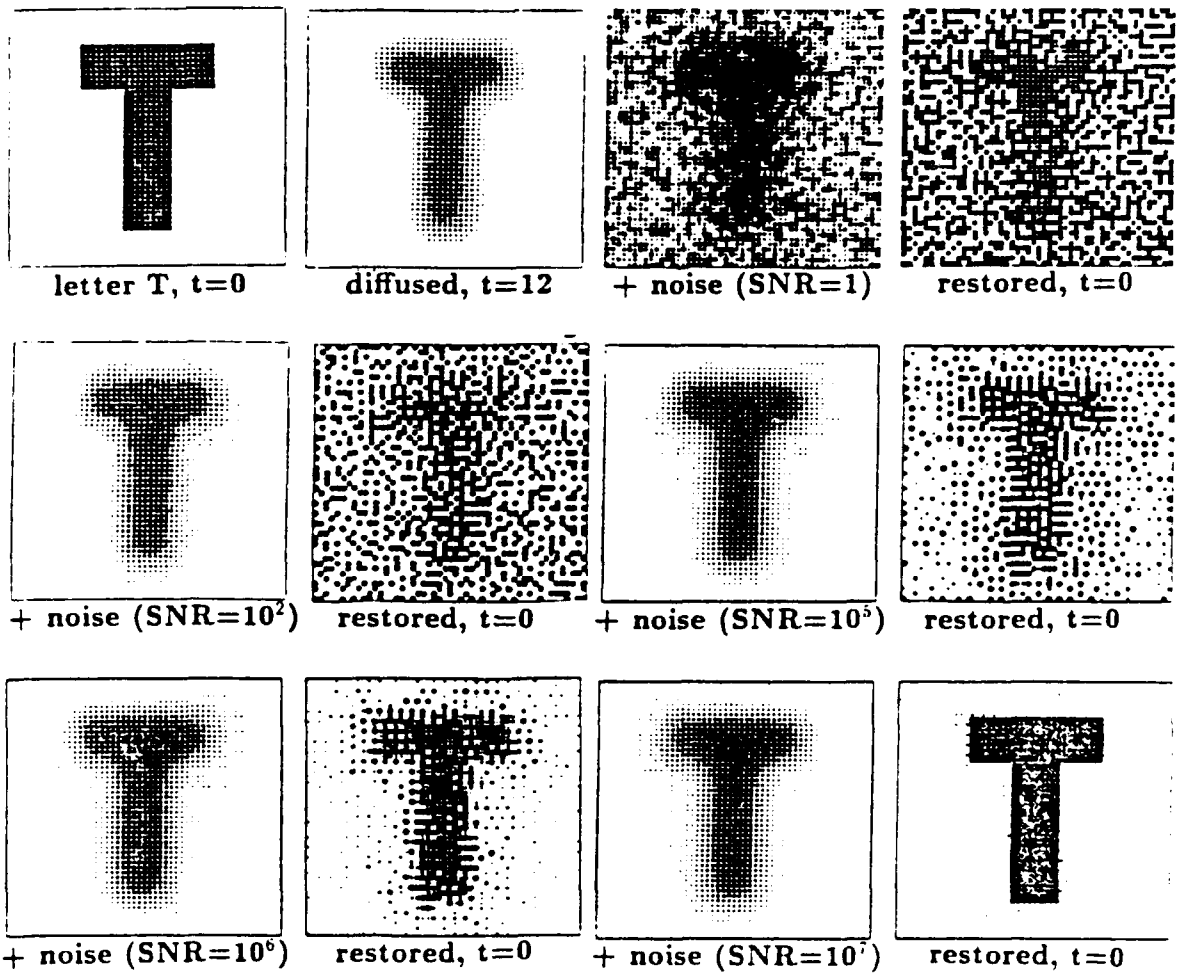
22. Letter "E" (for Entropy), which has been diffused for a duration of time $t = 24$ (12×10^4 iterations of the discretized diffusion equation), and subsequently reconstructed back to $t = 0$ (with intermediary results) by iterating the discretized antidiffusion equation (12×10^4 iterations). The reconstruction is successful.



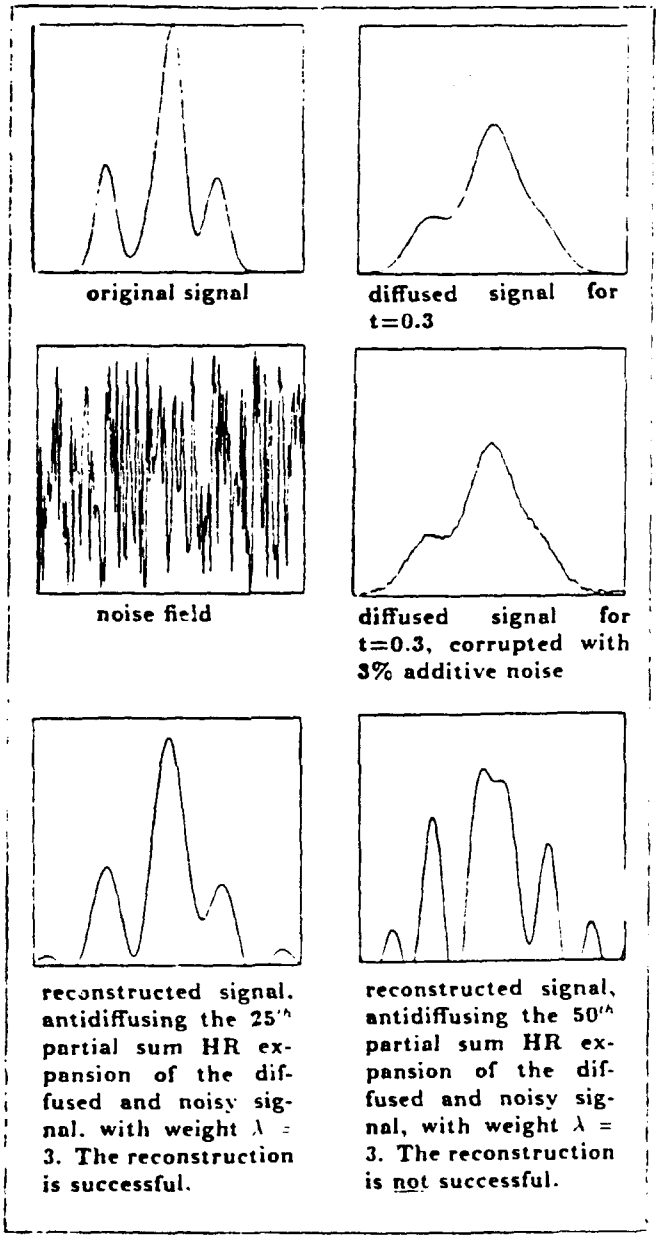
23. Letter "E", which has been diffused for a duration of time $t = 30$ (15×10^4 iterations of the discretized diffusion equation), and subsequently reconstructed back to $t = 0$ (with intermediary results) by iterating the discretized antidiffusion equation (15×10^4 iterations). The $t = 0$ reconstruction is not successful.



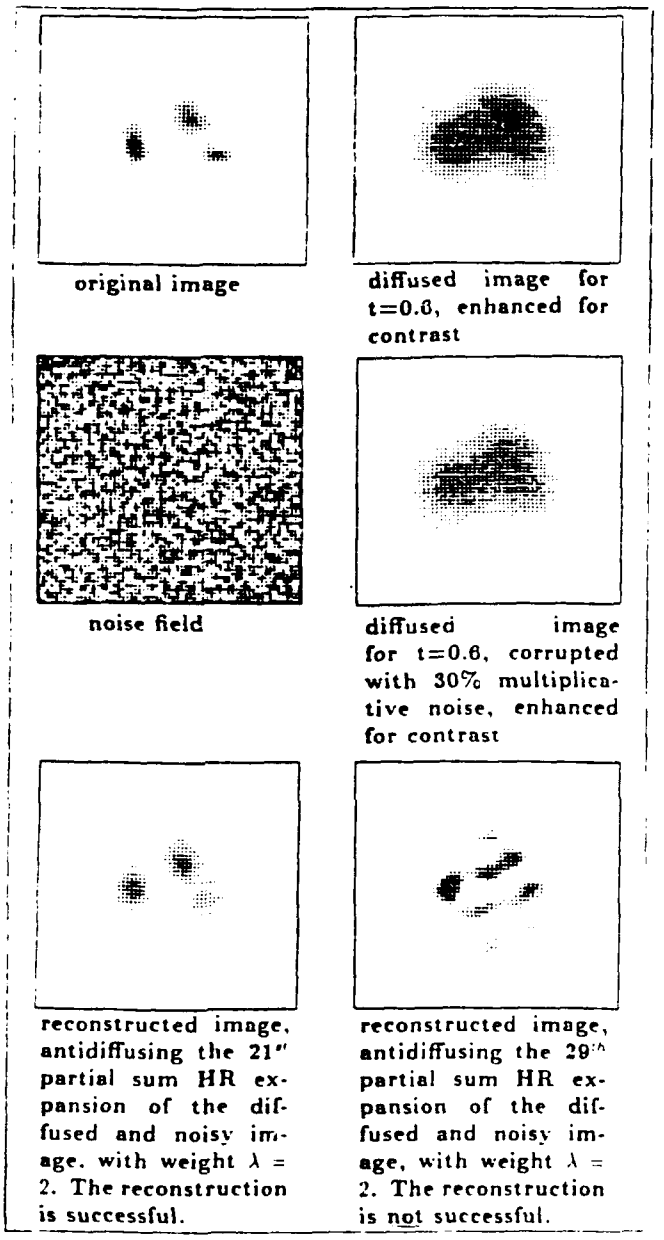
24. Spiral pattern, which is diffused for durations t=24, 25.8, 26.4, and 27, and subsequently reconstructed back to $t = 0$ by iterating the discretized antidiffusion equation. As time is increased, the reconstruction is increasingly unsuccessful.



25. Letter "T", which is diffused for t=12, and corrupted by additive noise at SNR's = 1, 10^2, 10^5, 10^6, and 10^7. The reconstruction by iterating the discretized antidiffusion equation is subsequently attempted. The reconstruction is only successful for extremely high SNR's.



26. Landscape 11, which is diffused for $t = 0.3$ by convolution with the Green's function of the 1+1 diffusion equation. The diffused signal is subsequently corrupted with 3% additive noise (the noise field is added uniformly onto the diffused signal, at a constant proportion of 3% of the maximum intensity of the signal). The reconstruction by antidiffusion of the HR wavelet expansion of the diffused and noisy signal is attempted. The weight of the HR wavelet expansion is $\lambda = 3$. The 25^{th} partial sum of the reconstruction is a good approximation to the original signal. The 50^{th} partial sum is not. The text explains why this happens.



original image

diffused image for t=0.6, enhanced for contrast

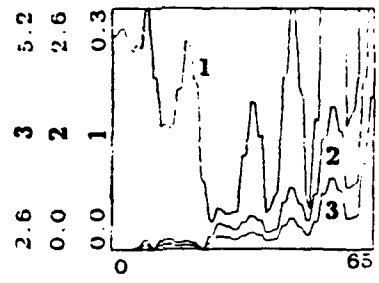
noise field

diffused image for t=0.6, corrupted with 30% multiplicative noise, enhanced for contrast

reconstructed image, antidiffusing the 21st partial sum HR expansion of the diffused and noisy image, with weight λ = 2. The reconstruction is successful.

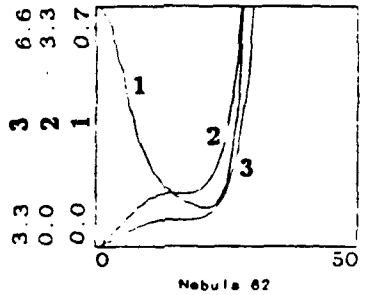
reconstructed image, antidiffusing the 29th partial sum HR expansion of the diffused and noisy image, with weight λ = 2. The reconstruction is not successful.

27. Nebula 62, which is diffused for $t = 0.6$ by convolution with the Green's function of the 2+1 diffusion equation. The diffused signal is subsequently corrupted with 30% multiplicative noise (the noise field is added pixel-by-pixel onto the diffused image, at a varying proportion of 30% of the intensity of the pixel). The reconstruction by antidiffusion of the HR wavelet expansion of the diffused and noisy image is attempted. The weights of the 2D HR wavelet expansion are $\lambda = \mu = 2$. The 21^{st} partial sum (ie. $\sum_n \sum_{m}^{n+m=21}$) of the reconstruction is a good approximation to the original image. The 29^{th} partial sum is not. The text explains why this happens.



Landscape 11

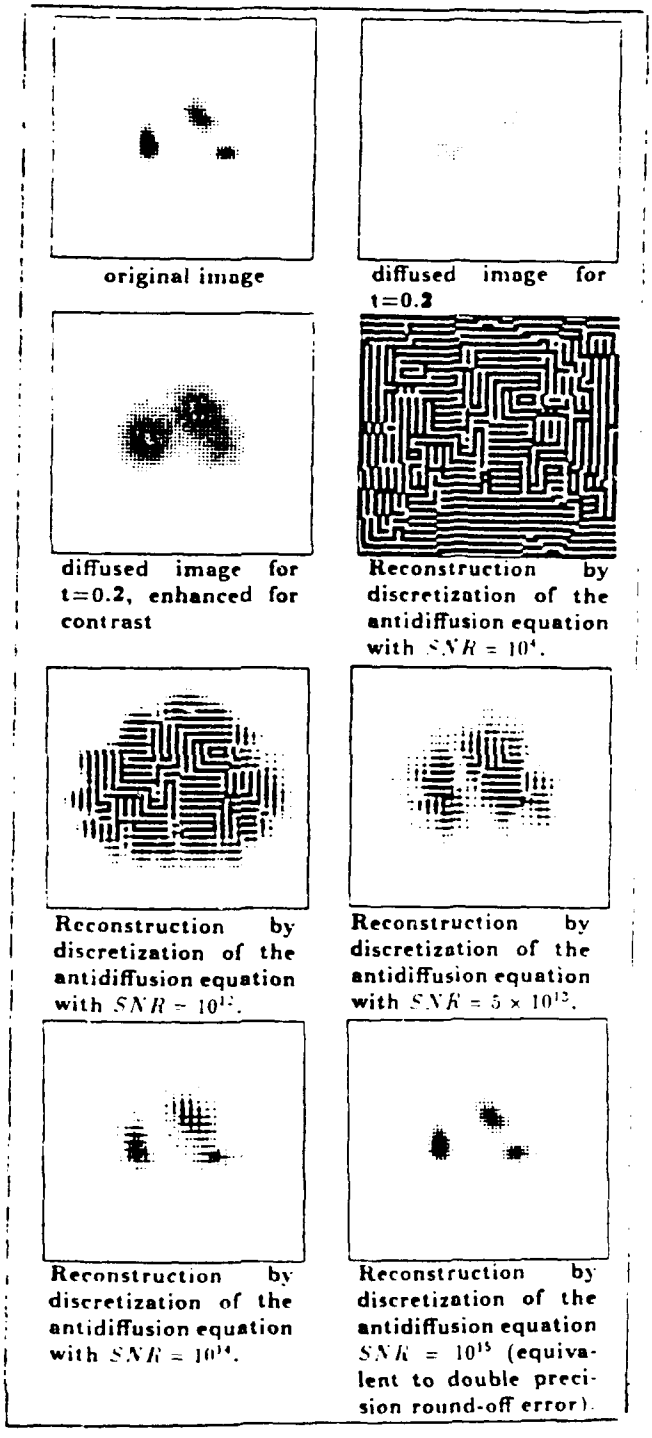
order	amount of negative intensity (curve 3)	total intensity (2.604766) (curve 2)	least squares fit (curve 1)
0	0.0000000E-00	2.604749	0.2650509
5	2.4976430E-03	2.608753	0.2497481
10	1.2247832E-02	2.627676	0.2135760
15	5.6897549E-02	2.716218	0.1514755
20	3.3983361E-02	2.673356	0.2249542
25	7.7245735E-02	2.7[illegible]	3.377[illegible]78E-02
30	0.1075234	2.822484	4.7781941E-02
35	0.1938762	2.995173	0.2123651
40	0.1120340	2.832430	7.3198855E-02
45	0.3814833	3.370775	0.3614376
50	0.3252646	3.259001	0.1912367
55	0.8707346	4.348985	0.9872774
60	0.3846842	3.377549	0.2134070
65	1.896428	6.399046	2.846997



Nebula 62

order	amount of negative intensity (curve 3)	total intensity (3.287609) (curve 2)	least squares fit (curve 1)
0	0.0000000E-00	3.287608	0.7088869
5	0.1504709	3.588553	0.5150415
10	0.3304867	3.948536	0.2440077
15	0.3568582	4.000921	0.1455814
20	0.3910236	4.069650	0.1138484
25	0.7805704	4.848684	0.1799982
30	3.332227	9.952042	2.049304
35	13.27546	29.83773	25.30508
40	47.56909	98.40572	304.2893
45	144.3484	291.9828	2652.888
50	395.5940	794.4598	19052.69

28. Positivity and total intensity conservation test for the reconstructions in the two previous figures. These tests yield the order of the best possible reconstruction by wavelet antidiffusion of blurred waveforms. As these tests fail (as curves 2 and 3 diverge), the reconstruction fails as well (the least-squares fit of the reconstruction with the original waveform, i.e. curve 1, diverges).



29. Attempt for reconstruction of the blurred image of nebula 62, by iterating the discretized antidiffusion equation. While HR wavelet antidiffusion allows reconstruction with very low SNR's, it can be seen that antidiffusion by discretization of the antidiffusion equation is only possible for very high SNR's.

Appendix A: Hyperdistribution Expansion of the Convolution Inverse of a Gaussian Filter

The Gaussian filter $\delta_\lambda(x,y)$ is a "radial" filter, that is, it is separable in its two independent variables x and y. As a result, its hyperdistribution expansion is equally separable, and the 2D Bochner Algebra reduces to the product of two 1D Algebras (see Eq.(165)). The hyperdistribution expansion of its inverse is thus equally separable, and We can restrict the study to just one variable, without loss of generality. The hyperdistribution expansion in x is obtained from Eq.(163), which is repeated here

$$\delta_\lambda(x) = \sum_{n=0}^{+\infty} (\frac{\lambda^4}{8})^n \frac{1}{n!} \nabla_x^{2n} \delta(x) \tag{210}$$

The 1D Bochner algebra thus reduces to determining the coefficients b_k such that

$$\sum_{k=0}^{n} \frac{(\frac{\lambda^4}{8})^{n-k}}{(n-k)!} b_k = \delta_{n0} \tag{211}$$

where $\delta_{n0} = 1$ if $n = 0$ and 0 otherwise. Trying the substitution

$$b_k = \frac{(-\frac{\lambda^4}{8})^k}{k!} \tag{212}$$

in the sum in Eq.(211) yields:

$$(\frac{\lambda^4}{8})^n \sum_{k=0}^{n} \frac{(-1)^k}{(n-k)!k!} \tag{213}$$

The partial sum above can be rewritten as

$$\sum_{k=0}^{n} \frac{(1)^{n-k}(-1)^k}{(n-k)!k!} \tag{214}$$

and the binomial expansion of $(1-1)^n/n!$ is recognized, which is identically equal to 0, except for when $n = 0$, in which case it is equal to 1. *q.e.d.*

Appendix B: Discretization of the operator ∇^2

In this appendix, we discretize the continuous operator ∇^2, in order to discretize both diffusion and antidiffusion equations. The discretization scheme will be central-in-space, and forward in time. Let $x_{00}, x_{01}, ..., x_{0n}, x_{10}, x_{11}, ..., x_{1n}, ...$ denote the spatial location of the grid points, and let the real function f, of the variable x_{pq}, be the measure of a real-valued distribution assigned to the grid. If confusion does not arise, we will denote $f(x_{pq})$ as f_{pq}. A Taylor series of order 2×2, for f_{ij}, around x_{pq} can be written as:

$$f_{ij} = \sum_{n=0}^{2} \sum_{m=0}^{2} \frac{1}{n!m!} (x_{iq} - x_{nq})^n (x_{jq} - x_{mq})^m \frac{\partial^{n+m}}{\partial n \partial m} f_{nm} + \mathcal{O}(\delta_x^3) \tag{215}$$

If $\{x_{ij}\}_{i,j}$ are the locations of all eight neighbors surrounding x_{pq}, and α_{ij} is the weight assigned to the neighbor at x_{ij}, we write:

$$\begin{aligned} \sum_{i=1}^{3} \sum_{j=1}^{3} \alpha_{ij} f_{ij} &= \sum_{i=1}^{3} \sum_{j=1}^{3} \alpha_{ij} \sum_{n=0}^{2} \sum_{m=0}^{2} \frac{1}{n!m!} (x_{iq} - x_{nq})^n (x_{jq} - x_{mq})^m \frac{\partial^{n+m}}{\partial n \partial m} f_{nm} + \mathcal{O}(\delta_x^3) \\ &= \sum_{n=0}^{2} \sum_{m=0}^{2} \left\{ \sum_{i=0}^{3} \sum_{j=0}^{3} \alpha_{ij} (x_{iq} - x_{nq})^n (x_{jq} - x_{mq})^m \right\} \frac{1}{n!m!} \frac{\partial^{n+m}}{\partial n \partial m} f_{nm} \\ &\quad + \mathcal{O}(\delta_x^3) \end{aligned} \tag{216}$$

We are interested in evaluating $\nabla^2 f_{pq}$ as a linear combination of the type $\sum_i \sum_j \alpha_{ij} f_{ij}$, where $\{x_{ij}\}_{i,j}$ are the locations of all eight neighbors of x_{pq}. Since

$$\nabla^2 f_{ij} = \frac{\partial^2}{\partial i^2} f_{ij} + \frac{\partial^2}{\partial j^2} f_{ij} \tag{217}$$

We set $\{\quad\}_{n,m} = \{\quad\}_{m,n} = 0$ for (m,n) in $\{(0,0), (0,1), (1,1)\}$, and $\{\quad\}_{n,m} = \{\quad\}_{m,n}$ for $(m,n) = \{(0,2)\}$, in Eq.(216). These constraints generate five equations, which reduce to the following two independent equations:

$$\nabla^2 f_{pq} = \frac{\alpha \sum f|_\diamond + \beta \sum f|_\square + \gamma f_{pq}}{(\alpha + \beta)\delta_x^2} + \mathcal{O}(\delta_x^4) \tag{218}$$

$$\gamma = -4(\alpha + \beta) \tag{219}$$

where α with no subscripts denotes the coefficient assigned to each cell of the diamond stencil, while β denotes the coefficient assigned to each cell of the square stencil, and γ is the coefficient assigned to the center-cell x_{pq}. α and β remain free parameters in the discretization.

Appendix C: Global Invariance of the Set of Parametric Hermite-Rodriguez Wavelets in the Diffusion Group

In this appendix, we show that the diffusion or antidiffusion of an HR wavelet yields a rescaled HR wavelet of same order, and of respectively bigger or lesser weight. Since the Dirac delta is the unit element for the operation of convolution, and since the diffusion of the Dirac delta yields a Gaussian of width two times the square root of the diffusion duration, we can write that:

$$e^{t\nabla^2} W_n^\lambda(x) = e^{t\nabla^2}(\delta(x) * W_n^\lambda(x)) = (e^{t\nabla^2}\delta(x)) * W_n^\lambda(x) = \frac{e^{-x^2/4t}}{2\sqrt{\pi t}} * W_n^\lambda(x) \quad (220)$$

The last operation of convolution above can be evaluated explicitly by employing the Rodriguez formula for HR wavelets, and the chain rule on ∇^n:

$$\frac{e^{-x^2/4t}}{2\sqrt{\pi t}} * W_n^\lambda(x) = \int_{-\infty}^{+\infty} \frac{e^{-y^2/4t}}{2\sqrt{\pi t}} \frac{(-1)^n \lambda^n \nabla_{x-y}^n}{\sqrt{2(n!)}} \frac{e^{-(x-y)^2/\lambda^2}}{\sqrt{\pi}\lambda} dy \quad (221)$$

$$= \frac{(-1)^n \lambda^n \nabla^n}{\sqrt{2(n!)}} \int_{-\infty}^{+\infty} \frac{e^{-y^2/4t}}{2\sqrt{\pi t}} \frac{e^{-(x-y)^2/\lambda^2}}{\sqrt{\pi}\lambda} dy \quad (222)$$

$$= \frac{(-1)^n \lambda^n \nabla^n}{\sqrt{2(n!)}} \frac{e^{\frac{x^2}{\lambda^2+4t}}}{\sqrt{\pi(\lambda^2+4t)}} \quad (223)$$

$$= \frac{\lambda^n}{\left(\sqrt{\lambda^2+4t}\right)^n} W_n^{\sqrt{\lambda^2+4t}}(x) \quad (224)$$

And consequently, we can write:

$$\text{diffusion:} \quad e^{+t\nabla^2} W_n^{\lambda_1}(x) = \left(\frac{\lambda_1}{\lambda_2}\right)^n W_n^{\lambda_2}(x) \quad (225)$$

$$\text{antidiffusion:} \quad e^{-t\nabla^2} W_n^{\lambda_2}(x) = \left(\frac{\lambda_2}{\lambda_1}\right)^n W_n^{\lambda_1}(x) \quad (226)$$

$$\text{with} \quad \lambda_2 = \sqrt{\lambda_1^2 + 4t} \quad (227)$$

In other words, the set of parametric HR wavelets is a globally invariant subspace -in the weight parameter λ- of the diffusion/antidiffusion group[14]. This property of HR wavelets enables one to formally diffuse/antidiffuse one-dimensional signals, or two-dimensional images, by diffusing/antidiffusing the HR wavelets, which are the building blocks of the signals' or images' HR expansions.

*Appendix D: Global Invariance of the Set of Parametric Hermite-Rodriguez Wavelets in the Convolution Group

This appendix is a generalization of the previous appendix: We show that the convolution of two HR wavelets of respective orders n and m, and respective weights λ and μ, yields another HR wavelet, of order $n+m$, and weight $\sqrt{\lambda^2+\mu^2}$.

Lemma 1. The Rodriguez formula for Hermite polynomials yields:

$$(-1)^n H_n(x)e^{-x^2} = \nabla^n(e^{-x^2}) \tag{228}$$

It can be rescaled to yield:

$$(-1)^n H_n(x/\lambda)e^{-x^2/\lambda^2} = \nabla^n(e^{-x^2/\lambda^2}) \quad \square \tag{229}$$

Lemma 2. By definition of HR wavelets,

$$W_n^\lambda(x) = Hd_n^\lambda(x) = \frac{H_n(x/\lambda)}{2^{n/2}\sqrt{n!}} \frac{e^{-x^2/\lambda^2}}{\sqrt{\pi}\lambda} \tag{230}$$

By Lemma 1,

$$W_n^\lambda(x) = \frac{(-1)^n\lambda^n}{2^{n/2}\sqrt{n!}} \nabla^n(\frac{e^{-x^2/\lambda^2}}{\sqrt{\pi}\lambda}) \tag{231}$$

and thus

$$W_n^\lambda(x) * W_m^\mu(x) = \frac{(-1)^{n+m}\lambda^n\mu^m}{2^{\frac{n+m}{2}}\sqrt{n!m!}} \nabla^n(\frac{e^{-x^2/\lambda^2}}{\sqrt{\pi}\lambda}) * \nabla^m(\frac{e^{-x^2/\mu^2}}{\sqrt{\pi}\mu}) \quad \square \tag{232}$$

Lemma 3. The convolution of the n^{th} derivative of e^{-x^2/λ^2} with the m^{th} derivative of e^{-x^2/μ^2} is given by

$$\nabla^n(e^{-x^2/\lambda^2}) * \nabla^m(e^{-x^2/\mu^2}) = \int_{-\infty}^{+\infty} \nabla_y^n(e^{-y^2/\lambda^2}) \nabla_{x-y}^m(e^{-(x-y)^2/\mu^2})dy \tag{233}$$

By the chain rule on ∇_{x-y}^m,

$$\begin{aligned} \nabla^n(e^{-x^2/\lambda^2}) * \nabla^m(e^{-x^2/\mu^2}) &= \int_{-\infty}^{+\infty} \nabla_y^n(e^{-y^2/\lambda^2}) \nabla_x^m(e^{-(x-y)^2/\mu^2})dy \\ &= \nabla^m \int_{-\infty}^{+\infty} \nabla_y^n(e^{-y^2/\lambda^2}) e^{-(x-y)^2/\mu^2} dy \end{aligned} \tag{234}$$

By a change of variables, and the chain rule on ∇_{x-y}^n,

$$\nabla^n(e^{-x^2/\lambda^2}) * \nabla^m(e^{-x^2/\mu^2}) = \nabla^{n+m} \int_{-\infty}^{+\infty} e^{-y^2/\lambda^2} e^{-(x-y)^2/\mu^2} dy \tag{235}$$

and finally,

$$\nabla^n(e^{-x^2/\lambda^2}) * \nabla^m(e^{-x^2/\mu^2}) = \sqrt{\frac{\pi}{\lambda^2+\mu^2}}\,\nabla^{n+m}(e^{-x^2/(\lambda^2+\mu^2)}) \quad \square \tag{236}$$

Finally combining Lemmas 2 and 3,

$$W_n^{\lambda}(x) * W_m^{\mu}(x) = \gamma_{n,m}^{\lambda,\mu}\, W_{n+m}^{\sqrt{\lambda^2+\mu^2}}(x) \tag{237}$$

with

$$\gamma_{n,m}^{\lambda,\mu} = \frac{\lambda^n \mu^m}{(\lambda^2+\mu^2)^{\frac{n+m}{2}}} \sqrt{\frac{(n+m)!}{n!m!}} \tag{238}$$

In other words, the set of parametric HR wavelets is a globally invariant subspace –in the weight parameter λ- under convolution. This property of HR wavelets enables one to formally convolve/deconvolve one-dimensional signals, or two-dimensional images, by convolving/deconvolving the HR wavelets, which are the building blocks of the signals' or images' HR expansions.

Appendix E: Waveform Total Intensity Conservation in the Diffusion Group

We first prove that any diffusion/antidiffusion process conserves the total intensity of the initial condition. In other words, if We denote by $I(x)$ the initial condition of a diffusion/antidiffusion process, then the total intensity is given by $\int_{-\infty}^{+\infty} I(x)\,dx$, and

$$\frac{\partial}{\partial t} \int_{-\infty}^{+\infty} e^{\pm t\nabla^2} I(x)\,dx = 0 \tag{239}$$

The proof is as follows. By integrating in space the diffusion/antidiffusion equation for $u(x,t)$:

$$\int_{-\infty}^{+\infty} \frac{\partial u}{\partial t}\,dx \pm \int_{-\infty}^{+\infty} \nabla^2 u\,dx = 0 \tag{240}$$

With the help of Gauss' theorem,

$$\frac{\partial}{\partial t} \int_{-\infty}^{+\infty} u\,dx \pm \left[\vec{n}\cdot\vec{\nabla}u\right]_{-\infty}^{+\infty} = 0 \tag{241}$$

and thus

$$\frac{\partial}{\partial t} \int_{-\infty}^{+\infty} u\,dx = 0 \tag{242}$$

Since $u(x,t) = e^{\pm t\nabla^2} I(x)$, we obtain the desired result:

$$\frac{\partial}{\partial t} \int_{-\infty}^{+\infty} e^{\pm t\nabla^2} I(x)\,dx = 0 \tag{243}$$

The same reasoning can be easily extended to the case where the initial condition is a two-dimensional image $I(x,y)$.

Furthermore, the total intensity of a truncated HR wavelet expansion of any waveform is exactly equal to the total intensity of the waveform. That is because the total intensities of all HR wavelets are identically equal to zero, except for the total intensity of the HR wavelet of order 0, which is by definition the zeroth HR moment, and thus the total intensity of the waveform. In other words, the expansion of a waveform in terms of a truncated HR wavelet series conserves the total intensity of the original waveform (this property is also true in Fourier series).

*Appendix F: Do Point-Spread-Functions Conserve the Total Intensity of the Original Waveform?

In this appendix, we derive the constraint imposed on the point-spread function of LTI systems which conserves the total intensity $\mathcal{I}$ of their input. In other words, if $u(x)$ denotes the input function, $v(x)$ denotes the output, and $f(x)$ denoted the impulse response or point-spread-function, we require that

$$\int_{-\infty}^{+\infty} v(x)dx = \int_{-\infty}^{+\infty} u(x)dx = \mathcal{I} \tag{244}$$

But

$$v(x) = u(x) * f(x) = \int_{-\infty}^{+\infty} f(x-x')u(x')dx' = \int_{-\infty}^{+\infty} f(x')u(x-x')dx' \tag{245}$$

and thus, by exchanging the order of integration,

$$\begin{aligned} \int_{-\infty}^{+\infty} v(x)dx &= \iint_{-\infty}^{+\infty} f(x')u(x-x')dx'dx = \int_{-\infty}^{+\infty} f(x')\left\{\int_{-\infty}^{+\infty} u(x-x')dx\right\}dx' \\ &= \int_{-\infty}^{+\infty} f(x')\,\mathcal{I}\,dx' \end{aligned} \tag{246}$$

We conclude that the sufficient condition for total intensity conservation of the input is that the total intensity of the point-spread-function be normalized to

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \tag{247}$$

References

[1] A. Oppenheim, A. Willsky with I. Young, *Signals and Systems*, Englewood Cliffs, Prentice-Hall, N.J. (1978).

[2] L. Andrews, *Special Functions for Engineers and Applied Mathematics*, Macmillan Publishing Company, N.Y. (1985).

[3] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover Publications Inc., N.Y. (1970).

[4] L. Schwartz, *Théorie des distributions*, Hermann et Cie., Paris. (1951).

[5] I. Halperin, *Introduction to the theory of distributions* Canadian Mathematical Congress, University of Toronto Press, Toronto (1952).

[6] G. Temple, *Generalized Functions.* Proc. Roy. Soc. A, 228, 175-90 (1955), and J. Lond. Math. Soc. **28**, 134-48 (1953).

[7] J.P. Marchand, *Distributions, p.13.* North Holland Publishing Company, Amsterdam (1962).

[8] A.Erdelyi, *Operational Calculus and Generalized Functions*, Holt, Rinehart, and Winston, N.Y. (1962).

[9] I.M. Gelfand et al, *Generalized Functions*, Academic Press, N.Y. (1964-68).

[10] M.J. Lighthill, *Fourier Analysis and Generalized Functions*, Cambridge University Press, Cambridge (1970).

[11] F.G. Friedlander, *Introduction to the theory of distributions* . Cambridge University Press, N.Y. (1982).

[12] S. Bochner and W.Martin, *Several Complex Variables*, Princeton University Press, Princeton N.J. (1948)

[13] K. Ito, Ed., *Encyclopedic Dictionary of Mathematics, 2nd Ed*, MIT Press, Cambridge, MA (1987).

[14] C. Konstantopoulos, L. Mittag, G.H. Sandri and R. Beland, *Deconvolution of Gaussian Filters and Antidiffusion*, J. Appl. Phys. pp.1415-20, Aug 1, (1990).

[15] B.A. Rhodes, *Integral Equations and Inverse Problems*, Ph.D. Thesis, Boston University Graduate School, Division of Engineering and Applied Science, Boston MA. 02215 (1991).

[16] J.G. Mikusiński, *Fundam. Math.* **35** (1948).

[17] J.G. Mikusinski, *Operational Calculus*, Pergamon Press, NY (1959).

[18] J.D. Bjorken and S.D. Drell, *Relativistic Quantum Mechanics*, Mc-Graw Hill, N.Y. (1964).

[19] C. Konstantopoulos, and G. Sandri, *Schrödinger Free-Particle Flow Group, and the Classical Diffusion Semi-Group*, Il Nuovo Cimento, accepted for publication, March 1, 1991.

[20] E. Hille and R. Philips, *Functional Analysis and Semi-Groups*, American Mathematical Society, Waverly Press, Baltimore, MD. (1957).

[21] P. Morse, *Thermal Physics*, Second Edition, W.A. Benjamin , N.Y. (1969).

[22] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill, N.Y. (1953).

[23] G. Wentzel, *Quantum Theory of Fields. Translation of Einfuhrung in die Quantentheorie der Wellenfelder. p.177*, Interscience Publishers, N.Y. (1949).

[24] M. Greenberg, *Applications of Green 's Function in Science and Engineering*, Prentice-Hall, Inc. N.J. (1971).

[25] T.M. Apostol, *Mathematical Analysis*, Addison-Wesley, Reading MA. (1960).

[26] L. Hormander, *The Analysis of Linear Partial Differential Operators, vol.1, p.181, Paley-Wiener-Schwartz theorem*, Springer-Verlag, Berlin , N.Y. (1983).

[27] Y. Meyer, *Wavelets and Operators*, CEREMADE, Univérsité de Paris-Dauphine, no.8704 (1989).

[28] R. Vichnevetsky and J.B. Bowles, *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*, SIAM, Philadelphia PA. (1982).

[29] F. Riesz and B. Sz-Nagy, Translation of *Leçons d'Analyse Fonctionelle*, N.Y. Unger, (1955).